

A COMPUTATIONAL MODEL OF “ARTIFICIAL INTUITION” IN  
DECISION MAKING

BY

OLAYINKA PATRICK JOHNNY

A THESIS SUBMITTED  
IN PARTIAL FULFILMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

EDGE HILL UNIVERSITY  
ORMSKIRK, ENGLAND  
SEPTEMBER 2020

# **A Computational Model of "Artificial Intuition" in Decision Making**

**Olayinka Patrick Johnny**

Submitted for The Degree of  
Doctor of Philosophy

## **ABSTRACT**

The ability to perform a data-driven decision-making approach is at the core of Data Science, AI, and general Machine Learning techniques. To achieve a detailed data-driven approach, all possible scenarios must be considered, and their outcomes must be assessed logically and systematically to obtain accurate and applicable methods for knowledge discovery. These are considered in order to identify the best possible choice. Although, the data-driven approaches have been shown to be effective in theory, a major drawback is that it is typically associated with high computational complexity. Moreover, it is non-trivial to develop and train models with deep and complex model structures with potentially large number of parameters. However, there are compelling evidence from the cognitive sciences that intuition plays an important role in intelligence extraction and the associated decision-making process. More specifically, intuition can be used to identify, combine and discover knowledge in a 'parallel' manner and so more efficiently. As a consequence, the embedding of Artificial Intuition within Data Science is likely to provide novel ways to identify and process information.

The first contribution of this thesis is the introduction of a rigorous mathematical formulations and a novel algorithm for artificial intuition. Specifically, a mathematical formulation is introduced to describe a model that utilises semantic network to improve decision making. Moreover, the model that is introduced included some lemmas and propositions that provides a way of combining the aggregation of edges in semantic networks. The model implements techniques from computational linguistic via the processing pipeline to derive semantic networks. Moreover, the thesis contributed a state-of-the-art review of artificial intuition. The author provided relevant

and detailed research about the concepts of artificial intuition as it relates to creativity, gut-feeling, rational thinking. It finally identified the umbrella concept called artificial intuition and identified some key requirements for the development of a model.

**Keywords:** Artificial Intuition, Artificial Intelligence, Modelling, Network Theory, Data Science, Decision Algorithms

## **Dedications**

I will like to dedicate this Ph.D thesis to my wife and children for their love, support and encouragement. I will also like to dedicate this to my parents, even though they are no longer here with us, they will be proud of their son.

## Acknowledgements

I will like to thank all those who have made this thesis possible.

I will like to thank my Ph.D supervisory team, Dr. Swagat Kumar and Dr. Hari Pandey. In particular, my special thanks go to my Director of Studies, Prof. Marcello Trovati for his guidance and support and for supervising my thesis. I have learnt a lot from you in this my Ph.D journey. I will not forget to thank Professor Nik Bessis for his support.

I will like to thank my wife and angel, Toyin for her patience, support and understanding throughout my studies. My thanks also go to my wonderful children, Oluwadamilare and Oluwademilade for being so lovely.

## List of Main Publications

1. Johnny, O., Trovati, M., and Ray, J. (2017, August). The “Gut-Feeling” in the Decision-Making Process: A Computationally Efficient Approach to Influence Relation Assessment. In International Conference on Intelligent Networking and Collaborative Systems (pp. 576-584). Springer, Cham.
2. Johnny, O., Trovati, M., and Ray, J. (2019, August). Towards a Computational Model of Artificial Intuition and Decision Making. In International Conference on Intelligent Networking and Collaborative Systems. Springer, Cham.
3. Johnny, O., and Trovati, M. (2020, August). Knowledge-Based Networks for Artificial Intuition. In International Conference on Intelligent Networking and Collaborative Systems (pp. 319-325). Springer, Cham.
4. Johnny, O., and Trovati, M. (2018, September). Big Data Inconsistencies: A Literature Review. In International Conference on Intelligent Networking and Collaborative Systems (pp. 486-492). Springer, Cham.
5. Ray, J., Johnny, O., Trovati, M., Sotiriadis, S., and Bessis, N. (2018). The Rise of Big Data Science: A Survey of Techniques, Methods and Approaches in the Field of Natural Language Processing and Network Theory. *Big Data and Cognitive Computing*, 2(3), 22.
6. Johnny, O., and Trovati M. (Jan 2020). Big data inconsistencies and incompleteness: a literature review *Int. J. Grid and Utility Computing*,
7. Johnny, O., Sotiriadis, S., Asimakopoulou, E., and Bessis, N. (2014, September). Towards a social graph approach for modeling risks in big data and Internet of Things (IoT). In *Intelligent Networking and Collaborative Systems (INCoS)*, 2014 International Conference on (pp. 439-444). IEEE.

## Contents

<b>1</b>	<b>Introduction</b>	<b>12</b>
1.1	Problem Statement . . . . .	13
1.2	Research Aim and Objectives . . . . .	16
1.3	Motivation . . . . .	16
1.4	Major Contributions . . . . .	19
1.5	Thesis Organization . . . . .	20
<b>2</b>	<b>Background and Literature Review</b>	<b>22</b>
2.1	Background . . . . .	22
2.2	Human Intuition and Decision Making Approaches . . . . .	23
2.3	Overview of the application of artificial intuition . . . . .	30
2.3.1	Intuition and Clinical Decision Making (CDM) . . . . .	30
2.3.2	Intuition and Business Decision . . . . .	31
2.3.3	Intuition and Player in a Game . . . . .	35
2.4	Emotion and Decision Making Approaches . . . . .	35
2.5	Rational Decision-Making Approaches . . . . .	42
2.5.1	Bayesian Network Creation from data . . . . .	43
2.6	Guts-feeling and Decision Making . . . . .	50
2.7	Shallow Learning Approaches of Decision Network . . . . .	54
2.8	Neural Network and Decision-Making Approaches . . . . .	55
2.9	Summary and Discussion of the Approaches . . . . .	59
2.10	Key Requirements for the Model Development . . . . .	62
<b>3</b>	<b>Network, Graph and Semantic Network</b>	<b>66</b>
3.1	Network and Graph Theory . . . . .	66
3.1.1	Scale-free Networks . . . . .	67
3.1.2	Random Graphs . . . . .	68
3.1.3	Small-world Networks . . . . .	68
3.2	Graph Theory . . . . .	69
3.3	Semantic Analysis . . . . .	71
3.4	Semantic Network . . . . .	76

3.4.1	Advantages and Disadvantages of Semantic Network	80
3.4.2	Applications of Semantic Networks . . . . .	83
3.5	Measurements of Semantic Similarities and Relatedness .	84
3.5.1	A Motivating Discussion . . . . .	86
3.6	Spreading Activation . . . . .	89
3.7	Modelling Semantic Network evolution and Decision Mak- ing . . . . .	92
3.8	Algorithms For The Evolution of Semantic Networks . . .	93
3.9	Summary . . . . .	96
<b>4</b>	<b>Gut-feeling in Decision Making: An initial Ap- proach for Influence Relation Assessment</b>	<b>98</b>
4.1	Description of the Approach . . . . .	100
4.2	Summary . . . . .	105
<b>5</b>	<b>The Proposed Solution</b>	<b>106</b>
5.1	Defining Essential Components of the Model . . . . .	106
5.2	General Definitions and Background of Main model . . .	108
5.3	Description of the Main Artificial Intuition Model . . . .	112
5.3.1	Combination of Edge Attributes . . . . .	114
5.3.2	Explicit Calculation of the Node Combinations . .	117
5.4	The Assessment of Potential Solutions . . . . .	118
5.4.1	Intuition Index . . . . .	119
5.4.2	Propagation Index . . . . .	119
5.4.3	Edge Entropy . . . . .	120
5.5	Entropy of the Query Space . . . . .	122
5.5.1	Geometric Interpretation of the Query Space . . . . .	124
5.6	Summary . . . . .	125
<b>6</b>	<b>Model Evaluation and Results</b>	<b>126</b>
6.1	Description of Datasets: ConceptNet, Wikipedia and Recipe- Ingredient Datasets . . . . .	127



6.1.1	ConceptNet . . . . .	127
6.1.2	Wikipedia . . . . .	130
6.1.3	Recipe-Ingredient Dataset . . . . .	131
6.1.4	Validation Approach . . . . .	132
6.2	Evaluation and Results . . . . .	134
6.2.1	Evaluation using Recipe-ingredient and Concept- Net Dataset . . . . .	135
6.3	Summary . . . . .	143

## 7 Conclusion and Future Works 144

7.1	Limitation of the Research . . . . .	145
7.2	Future Work . . . . .	146

### List of Figures

1	Making a decision to take the umbrella . . . . .	15
2	Overview of entrepreneur decision making process . . . . .	34
3	A network graph from ConceptNet that is formed from the concept <i>Thinking</i> and the related edges (Liu and Singh, 2004) . . . . .	70
4	Graph Relationship between two Nodes extracted from texts in Wikipedia . . . . .	72
5	Example of entities and relations extracted from texts in Wikipedia . . . . .	74
6	Example 2 of Entities and Relations Extracted . . . . .	75
7	Example of “is-a” link that suggest ambiguity . . . . .	82
8	Example of “is-a” link after removing the ambiguity . . . . .	82
9	Example of “is-a” link that suggest self-loop . . . . .	82
10	Evolution of Semantic Network using algorithm 1. This is an example of a small-world network . . . . .	95
11	Influence Relation/Link in a Network . . . . .	101
12	The Components of Gut-feeling Method . . . . .	101
13	Network of Influence link between some <i>prior knowledge</i> of the dynamics of the rain . . . . .	102

14	Knowledge of the rain at a particular time in the past. It captures the reliability of the influence link over a period of time. This measures the entropy and the influence link in making a decision in the network . . . . .	103
15	Modelling Curiosity . . . . .	107
16	The Main Components of the Method . . . . .	108
17	The example discussed in Section 5.2. . . . .	110
18	A general depiction of artificial intuition process as the ‘light-bulb’ moment. . . . .	111
19	Overview of general components of the model . . . . .	113
20	Two simple networks as discussed in Section 5.3.1. . . . .	114
21	The “wings–propulsion–fly” network. . . . .	115
22	The “wheels–engine–noise” network. . . . .	116
23	The “dog–fish–animal” network. . . . .	116
24	A simple example. . . . .	117
25	Different highlighted paths in a small network. . . . .	123
26	Example of some semantic relations from ConceptNet (Liu and Singh, 2004) . . . . .	128
27	Semantic Network of a Car and Related Concepts (Liu and Singh, 2004) . . . . .	129
28	The components in developing the Artificial Intuition Model .	133
29	Graph of the weighted paths of ConceptNet . . . . .	138
30	Subnet based on the paths between FoodNet and ConceptNet with depth of 4 . . . . .	140
31	Paths from isi_ewu to Mexican Cuisine (Latin American) . .	141
32	Paths from isi_ewu to Brazilian Cuisine (Latin American) . .	141
33	The values of the entropy of the query solution for the validations described in Section 6.2 . . . . .	142

## List of Tables

1	Decision Options and Values . . . . .	33
2	Semantic Relation Ontology . . . . .	79

3	Sample raw data of the recipe ingredient dataset . . . . .	132
4	The validation results as discussed in Section 6.2 . . . . .	135
5	Co-occurrence of ingredients in FoodNet . . . . .	142

## List of Algorithms

1	Encode Semantic from Commonsense knowledgebase . . . . .	94
2	Mathematical Algorithm to define categories and find objects . . . . .	95
3	Ranking concepts and finding activation value . . . . .	96
4	Solution Assessment . . . . .	124

## 1. Introduction

This work focuses on the investigation and a rigorous modelling of artificial intuition in specific decision making scenarios. A major challenge in decision making is coping with a large unstructured data-sets based on partial unknown knowledge, and with multiple states which can be modelled by complex networks such as dependency network and Bayesian network. Making decision from such networks would require using the rational decision-making process which involve the need for deeper analysis to allow the investigation of their corresponding scenarios. Most of these models are logic driven and are time dependent (Dundas and Chik, 2011). Moreover, there are potentially a large number of parameters, which increase the overall computational complexity in some scenarios (Trovati and Bessis, 2016). The complexity relates to the amount of resources required to run the logic driven models for decision problems, in particular the time, storage and memory requirements are enormously high (Arora and Barak, 2009).

However, during critical and time based decision making scenarios such as medical diagnosis in an emergency, humans use intuition to solve problems (Payne et al., 2015; Graber et al., 2012; Hams, 2000). In this approach the author is particularly interested in the connection between artificial intuition and decision making and to develop a computational based model of artificial intuition in decision making. The research question is how Artificial Intuition can be developed and applied to specific decision scenarios. The theoretical underpinning of this study is drawn from the study of cognitive research. Human intuition is an unconscious mental process aimed to solve problems without using a rational decision-making process. Loosely speaking, intuition is the ability to see connections between concepts or properties that can be interpreted as valid with respect to some networks. That is by using intuitive models, a system is able to take subsets from two networks and pass them through a process to determine relationship that can be used to predict future decision without a deep understanding of a scenario and its corresponding parameters. When an artificial agent manifests human intuition properties, then we can say that there is Artificial

Intuition (Diaz-Hernandez and Gonzalez-Villela, 2017a). More specifically, Artificial Intuition is an automatic process, which does not search rational alternatives, jumping to useful responses in a short period of time, and is mainly focused in providing responses without iterative search of solutions for a given problem (Diaz-Hernandez and Gonzalez-Villela, 2017a)

In this study, artificial creativity and intuition refer to different, yet overlapping concepts. The former is associated with the automated creation and design of artefacts, objects, ideas, etc. which are regarded as (primarily) creatively ‘beautiful’. They could, of course, be useful and have various applications. On the other hand, artificial intuition involves the automated identification of innovative solutions. The main difference does not only focus on the above points. One of the main approaches used in artificial creativity involves specific machine learning and neural network methods. In fact, identifying creating patterns widely accepted by practitioners can be utilised as the building blocks of new automated creativity. Artificial intuition is assumed to focus on the identification of innovative solutions. Therefore, the semantic properties of the corresponding scenarios need to be investigated, by considering the semantic relationships and their corresponding dynamical properties need to be assessed to fully evaluate the suitability of a potential solution.

In this research study, a rigorous approach to artificial intuition is proposed. The aim is to facilitate a comprehensive theory and subsequent implementation of Artificial Intuition to provide a better decision system, which mimics the agile and efficient intuitive processes extensively used by human agents.

### *1.1. Problem Statement*

The research problem addressed in this thesis is the development of a computational model of artificial intuition in decision making. More specifically, it reflects the influence of intuition on decision-making process in critical situations. For example, consider the decision to take the umbrella when the weather is cloudy, and you want to go and see your favourite football game. In making such decision in a short time, the rational mind will require

some rational choices that are based on precise and logical approach, where as much information about the weather are checked. Whereas, an excited football fan will jump to the decision based on some shallow information which maybe far from being precise. Another example is to consider the case of a dealer in the business of buying and selling of stocks. How might the dealer’s intuition affect his decisions, and how might intuition help or hinder his planning process under critical situations and tight time constraints? A dealer in a positive emotional state might be more optimistic than statistically realistic estimates suggest and would not feel the need to look into the current market data. This is particularly so as the dependence on intuitive cues for decision-making during urgent situation becomes stronger, leading to less “rational” choices (Finucane et al, 2000). In fact, human intuition is the predominant disposition when human is under pressure and time restriction, therefore intuition kicks in and takes the important lead in driving the decision making process. In this context, knowledge based on historical data, personal experience and contextual knowledge drive the decision-making process rather than the deep rational process of decision making.

In this study, the author define artificial intuition as:

*”The ability of a system to assess a problem context and use pattern recognition or properties from a dataset to choose a course of action or aid the decision process in an automatic manner. In such scenarios, the innovative solution is accomplished through recognition of significant patterns and properties that are made available by prior knowledge and experience, given the context of the problem. The key idea here is that the prior knowledge and experience are captured as a mental model. Therefore, if we are analysing a problem scenario and we find a similar situation, the mental model is assessed to determine the semantic relatedness and similarity with the current scenario. If the overall semantic relatedness is consistent with present scenarios,*

*then it is determined to be accurate and an innovative solution can be reached.”*

One of the examples of a use case that describes the idea discussed in this research is presented in Figure 1.

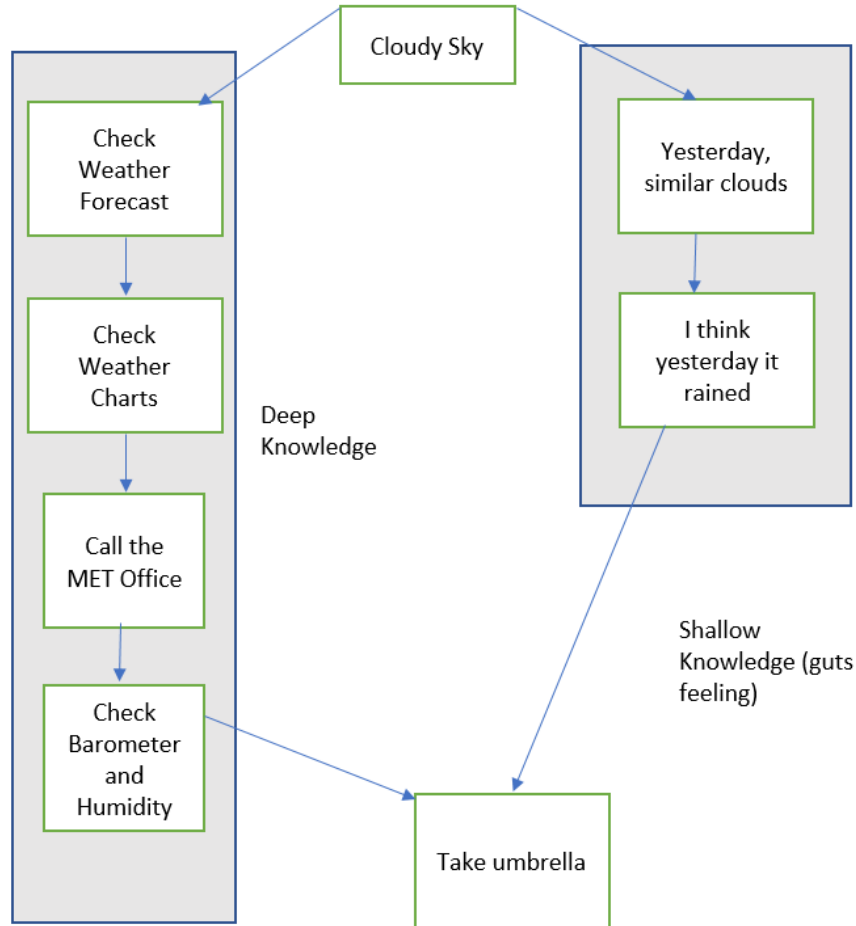


Figure 1: Making a decision to take the umbrella

Consider the decision whether to take the umbrella given that the sky is cloudy. On the left-hand side, the approach follows a more precise and logical approach, where it check as much information as possible. More specifically, there is the execution of the multiple normal processing of the

brain that follows the logical thinking approach and present itself as a choice to the user. Whereas on the right-hand side the approach ‘guess’ based on some shallow information, which maybe far from being precise! In other words, the fact that today the sky is as cloudy as it was yesterday, and that yesterday it (probably) rained, it is assumed to be enough to suggest that today it will also rain. Therefore, taking the umbrella is the right decision. This is basically the same as identifying a connection between weakly connected nodes in a network. Here, the evaluation of many nodes as possible are potentially bypassed to reach a satisfactory solution. Essentially, this decision is based on knowledge that is based on historical data, personal experience and expert knowledge.

### *1.2. Research Aim and Objectives*

This section presents a brief discussion and methodology of the research aim and the study objectives. The aim of this research is to study the connection between artificial intuition and decision making and to develop a computational based model of artificial intuition in decision making process. The driving question is how to utilise a model based on intuition to improve accuracy within a decision system

In view of this, the research study objectives are below:

1. Investigate and review current works on decision making approaches in large networks.
2. Identify key requirements via use-cases for intuition and decision-making process analysis.
3. Propose and develop a computational model based on artificial intuition and decision making.
4. To test and evaluate the model using suitably designed datasets.

### *1.3. Motivation*

A major challenge in decision making is coping with a large unstructured data-sets based on partial unknown knowledge, and with multiple states which can be modeled by complex influence diagrams, dependency network and Bayesian network. Making decision from such networks would



normally need a deeper analysis to allow the investigation of their corresponding scenarios. To this end, humans use logic in its best sense to perform a cost/benefit analysis that will provide the best possible choice. There are several algorithms to extract Bayesian Network from data, as well as to evaluate influence diagram (Suryadi and Gmytrasiewicz, 1999; Shachter and Bhattacharjya, 2010). These algorithms are based on a systematic rational thinking approach which assumes that humans are equipped with unlimited knowledge, time, and information-processing power. This also suggest a direct link between knowledge and the implementation of behavioral decisions, that is, that one does what one actually knows (Naqvi et al., 2006). Thus, in solving problems with such algorithms, a list of all different options and their possible outcomes are made, and then the use of logic in its best sense to perform a cost/benefit analysis that will provide us with the best possible choice must be identified and assessed (Velásquez, 1998). These assessments are often done independent of intuition.

However, when a decision process is carried out by expert modelers only partial amount of information is explicitly processed as they often rely on their intuition. There is compelling evidence from cognitive research that intuition play an important role in the rational process of intelligence and such decision-making process (Loewenstein and Lerner, 2003). More specifically, human intuition is the predominant disposition when human is under pressure and time restriction. In fact, intuition kicks in and takes the important lead in driving the decision making process. In this context, knowledge based on historical data, personal experience and expert knowledge drive the decision-making process rather than the deep rational process of decision making. More specifically, modeling of artificial intuition is accomplished through recognition of significant patterns and properties that are made available by prior knowledge and experience, given the context of the problem. Lieberman (2000) noted that with repeated exposures to certain problem domain and the pattern of solving such problems, the brain subconsciously learns patterns that form probabilistic relationships that are predictive of the significant events. In fact, when part of a previously-learned event occurs and the brain recognizes the elements that make up this event,

it anticipates and predicts subsequent elements resulting in a coherent overall interpretation of the partial data. Based on this recognition of pattern, the intuitive decision making kicks in.

Closely related to intuition is emotion. Emotions are changes in both body and brain states in response to different stimuli (Tranel and Damasio, 1991). Kumar et al. (2017) added that human emotions are mental states of feelings that arise spontaneously rather than through conscious effort and are accompanied by physiological changes in facial muscles which implies expressions on face. When physiological changes occur in the body, they are relayed to the brain where they are transformed into an emotion that tells the individual something about the stimulus that they have encountered. In fact, over time, these emotions and their corresponding states of the body become associated with particular situations and their past outcomes. Consequently, when making decision, these physiological changes and their evoked emotion are consciously or unconsciously associated with their past outcomes and therefore influence decision-making towards certain behaviours while avoiding others. This view of reasoning and decision-making is said to be associative because it compares similar situations that have been encountered in the past with current situations and then make decision accordingly. Thus, in doing this the intuition is invoked. On the other hand, some have argued that rational thinking and decision-making does not leave much room for intuition (Livet, 2010). Others have said that intuition is considered irrational occurrences that may distort reasoning (Barnes and Thagard, 1996). In Zeelenberg et al. (2008), the authors cautioned that we are only rational within the limits of our cognitive capacities and that decision making itself is often an emotional and intuitive process. Tranel and Damasio (1991) argue that without emotional involvement, decision making might not even be possible or might be far from optimal. In fact, decision making depends on emotional processing and the resulting intuitive feelings, which involve images that relate to the state of the body. This study agrees with the views of Zeelenberg et al. (2008) and Tranel and Damasio (1991). Forming the appropriate intuitions could help in decision making. Binali et al. (2010) sums it “reason without emotion and intuition is inad-

equate for making the decisions that guide our lives, and in fact make up our lives”. This provides motivation and a new perspective on incorporating intuition in decision making systems. However, effects of artificial intuition on decision-making have not been sufficiently addressed. Relatively little work has been done on creating computational models that incorporate the effect of intuition on the human capacity for decision-making (Langley et al., 2009; Gmytrasiewicz and Lisetti, 2001).

The concept of artificial intuition, as suggested by the early studies, focuses largely on the concept itself, rather than on the representation and use of entities in the process. While the earlier studies took a psychological and cognitive approach, some recent studies attempt to study artificial intuition from the computational point of view. However, an important limitation of these approaches is that they did not present a detailed representation of algorithm and use of intuitive entities in the process.

This research addresses this gap by developing a computational model of artificial intuition and decision making. Therefore, the author believe that a computational model based on intuition can potentially obtain accurate and optimal result and improve the overall performance of decision-making process. One potential real-world application of this model is in simulating weather prediction. For example, modeling the decision to take the umbrella when the weather is cloudy, and you want to go and see your favorite football game. Other potential application is in integration with Clinical Decision Support (CDS) system to reduce diagnostic errors; simulating the decision of a dealer in the business of buying and selling of stocks. In other words, predicting the accuracy of the buyers choice and simulating the decision of a player in a game,

#### *1.4. Major Contributions*

This research study presents the computational model for artificial intuition. This thesis has some major contribution as follows:

- First, it introduces a rigorous and a novel algorithm for artificial intuition. Specifically, a mathematical formulation is introduced to de-

scribe a model that utilises semantic network to improve decision making.

- It implements techniques from computational linguistic via the processing pipeline to derive semantic networks. Facts, ideas and memories in our brain are not structured into tables like in relational databases, rather the brain is a graph, network of interconnected nodes. Therefore, using semantic network-based model and representation of the data is a suitable representation of how the brain works.
- By using semantic network-based model, it is possible to apply spreading activation to simulate how human intuition work and quickly connect concepts thus improving the performance of intuitive decision making.
- Contributed a state-of-the-art review of artificial intuition. The author provided relevant and detailed research about the concepts of artificial intuition as it relates to creativity, gut-feeling, rational thinking. It finally identified the umbrella concept called artificial intuition and identified some key requirements for the development of a model.

### *1.5. Thesis Organization*

The thesis is structured as follows.

Chapter 2 presents a background and literature review of approaches related to decision making. Specifically, a state-of-the-art review of approaches in decision making in large networks is discussed as part of the study's objective 1. These approaches are discussed and evaluated in order to provide a compelling motivation for the development of the computational model of artificial intuition and decision making. Moreover, a set of essential requirements via use-cases for the development and implementation of artificial intuition and decision-making process analysis was developed.

The first part of Chapter 3 presents a discussion on data and knowledge inconsistencies and incompleteness and the implications for decision making. This discussion was presented with the aim to define completeness and stability of known knowledge over specific time periods as one of the objectives

of this research is to extract networks from structured and unstructured data sources. The second part of chapter 3 presents the theory and models of networks as the suitable approaches to be used in the definition and implementation of the model in this study. It provides a discussion of semantic networks since this constitute the representation model the author have chosen to implement the model. The chapter also presents a discussion of some semantic similarities and relatedness measures as well as approaches for identifying candidate terms for spreading activation in a network. The chapter concludes by presenting some algorithms for the modeling of the evolution of semantic networks creation and decision making.

Chapter 4 presents a description of a novel approach for assessing influence relationships between pairs of extracted concepts based on an interpretation of the *gut feeling*, which provides an accurate and computationally efficient approach to the decision-making process. This approach is presented with the aim to define complex decisional networks by bypassing complex calculations by providing an innovative approach to their analysis.

Chapter 5 presents the proposed model of the main artificial intuition. The discussion encompasses a general description of the main components of the model as well as the various mathematical descriptions and discussions that implements the model.

Chapter 6 presents the evaluation of the model which utilises knowledge network representation of the knowledge base. The evaluation is based on scenarios captured by semantic networks. It discusses the design of semantic network based on suitable textual analysis of large datasets. This knowledge network is essential in designing an Artificial Intuition framework. The intuition is informed by any general knowledge, as well as more contextualised and ‘intuitive’ knowledge.

Chapter 7 summarises the contributions, conclusions and future works.

## 2. Background and Literature Review

The purpose of this chapter is to gather and understand existing literature and research that has been performed relating to artificial intuition and decision making and to ensure that the correct aspects are understood and to discover what relevant theory that can be applied. The chapter consists of a background on artificial intuition and decision making and the approaches to extract and evaluate relevant information to assess influence relations between concepts that are used in a decision system

### 2.1. Background

There are a variety of approaches to extract and evaluate relevant information to assess influence relations between concepts (Shachter and Bhattacharjya, 2010). All these methods are based on rational thinking, which assume unlimited knowledge, time, and information-processing power. In other words, all possible scenarios are considered, and their outcomes are assessed via a logical and systematic manner to identify the best possible choice. The ability to perform such evaluation based on large quantities of parameters, which define a specific scenario, is at the core of deep learning. However, there is compelling evidence from neuroscience research that emotions and intuition play an important role in the rational process of intelligence and the decision-making process (Loewenstein and Lerner, 2003).

The interplay between the rational mind and the intuitive mind was summed up by Hart (2009) as follows:

*"The analytic grasps and holds, while the intuitive opens and embraces; the analytic has purpose, while the intuitive plays; the analytic measures and calculates, while the intuitive appreciates; the analytic builds, cuts, and controls, while the intuitive remains open-ended, the analytic is contained and directed by the ego and the will, while the intuitive tends toward self-transcendence and arises spontaneously, the analytic is willful, while the intuitive is willing."*

Decision-making involves the selection of a course of action from among two or more possible alternatives in order to arrive at a solution for a given problem. Every decision making process produces a final choice. The output can be an action or an opinion of choice. Stanovich and West (2000) noted that humans engage in two strategies of decision-making involving both cognitive and intuitive processes. This have been the general consensus among researchers that have studies decision making. The cognitive process represents an analytical or rational approach, which is explicit and systematic in nature while the intuitive process, which is implicit and global, consists of establishing salience among pieces of information and recognizing patterns that produce coherence (Stanovich and West, 2000; Lieberman, 2000; Smith and DeCoster, 2000). Some have suggested that the cognitive and intuitive processes can operate independently or can be complementary to one each other (Smith and DeCoster, 2000). The research on decision making has a long history and has been a primary focus of research among psychologists and cognitive scientists (Redelmeier et al., 2001).

In the remaining part of this section, the author present a literature review of the approaches related to decision making process. The author have categorized the analysis of the process in: *a) Human Intuition and decision-making approaches b) Rational Decision-making approaches c) Emotion and decision making, d) Guts-feeling and decision making, e) Shallow learning approaches of Bayesian Network f) Neural Network and decision making.*

## *2.2. Human Intuition and Decision Making Approaches*

This section provides a review of intuitive decision making approaches.

Intuition and the role in decision making have been studied from the psychology, philosophy and cognitive research perspective. Simon (1987) presents a descriptive research on managerial decision making and problem solving where he provides insights into the nature of intuition. He noted that executives who attend to real-time information are actually developing their intuition and aided by intuition, they can react quickly and accurately to changing stimuli in their firm or its environment. Although the data

are limited, the executives who relied most heavily on real-time information were also most frequently described as being intuitive. Simon (1987) explained that intuition is dependent on past knowledge and experience for better recall of solutions to the given problems or normal logical process. Poria et al. (2014) suggest intuition as the process of making analogies between the current problem and the ones solved in the past to find a suitable solution. Minsky (2007) attributes this property of intuition to the so called ‘difference-engines’, which are agents which operates by recognizing differences between the current state and the desired state and acting to reduce each difference by invoking K-lines that turn on suitable solution methods. Minsky noted that this kind of thinking may be the essence of our supreme intelligence since in everyday life no two situations are ever the same and we have to perform this action continuously.

Cert and Wilcockson (1996) argued that although intuition may be understood as irrational process, but it has a rational basis. Intuitive thinking has certain essential features and it involves the use of sound, rational and relevant knowledge base in situations that through experience are so familiar that the person has learned how to recognise and act on appropriate patterns. In other words, the intuitive decision is executed with help of a rational database of previously acquired knowledge but hidden from conscious mind which the human use to recognise and act rapidly. Minsky (2007) noted that these patterns are simple pieces of information that help to create the intuitive decision.

Kahneman (2002) have identified two fundamental and distinct modes of decision making in humans. The theory distinguished between intuition and reasoning. Stanovich and West (2000) labelled them as *System 1* (intuition) and *System 2* (reasoning). *System 1* is an automatic, fast and often unconscious way of thinking. It is autonomous and efficient, requiring little energy or attention and is dependent on some known information. *System 2* is an effortful, slow and deliberately controlled way of thinking. While *System 1* is likely to be affected by recognition of patterns, *System 2* is based on rational choices where humans use logic in its best sense to perform a cost/benefit analysis that will provide the best possible choice.



Studies have shown that human beings do not have the natural ability to do two things at the same time, hence some have argued that it is important to move from *System 2* (consciousness/rationality/using logic) to *System 1* (subconsciousness, intuition). Moving from *System 2* to *System 1* requires the agent/human to acquire a lot of skills and experiences over a period. This is very useful during critical and time based decision making scenarios such as the weather forecast and medical diagnosis in an emergency (Payne et al., 2015; Graber et al., 2012; Hams, 2000). Studies indicate most clinical decisions are made using the fast, hardwired intuitive *System 1* approach that depends heavily on the inductive reasoning experience of the decision maker (Payne et al., 2015; Graber et al., 2012). They further emphasized that the experience and recognition skills of the decision maker will determine how well the presented information is interpreted. Most System 1-based decisions are correct, but also subject to bias-induced error in cases where atypical signs/symptoms present, or when a non-specific pattern is mistakenly associated with the wrong diagnosis. Essentially, they recognized that intuition is accomplished through recognition of significant information and it synthesizes immediate decisions or actions without the need of a rational process (Kahneman, 2002). He also acknowledged that this depends on the use of prior knowledge and experiences and the context of the problem.

Wilson and Schooler (1991) showed an experiment to determine whether cognitive decision-making does lead to better decisions. In one experiment they had participants rate their preferences for the taste of several fruit-flavoured jam and compared with experts ratings of the jams, they noted that too much cognitive effort actually lowers the quality of performance, instead, most people are often happier with intuitive decisions than those based on deliberate, rationale-based determinations. Klein (2017) has argued that there are other situations in which skilled decision makers do better when they trust their intuitions than when they engage in detailed analysis, in particular when experienced decision makers are working under pressure rarely need to choose between options because, in most cases, only a single option comes to mind. In fact, in the clinical setting, Langan-Fox and Shirley (2003) have noted that health care practitioners use intuition when

information is incomplete, and the course of action is ambiguous. Likewise, similar application of this type of intuitive reasoning by human is found in the weather forecast (Langan-Fox and Shirley, 2003).

Sonntag (2011) explained how a multimodal dialogue intuitive system can be implemented by using intuition as a recommendation system. Sonntag (2011) argued that the fact that humans adapt their dialogue behaviour over time according to their dialogue partners' knowledge, attitude, and competence suggest that the influence of intuition in this natural human communication behaviour might be positive. He noted that an environment, where an intuition model extends a sensory-based modelling of instincts can help to assess the significance of intuition in multimodal dialogue. However, this work did not present any formal description of the concept of intuition.

He and He (2008) described experiential learning theory as the process whereby knowledge is created through the transformation of experience. According to Tao and He (2009), experiential learning emphasizes the role that appropriate environments and experiences play in the learning process as the learner is directly in touch with the realities being studied rather than merely thinking about the encounter or studying the experience of others with such phenomena. They presented an artificial intuition learning model called Trusted Intuition Network (TIN) for programmers and users to collaborate. The most basic conclusion from Tao and He (2009) theory is that people do learn from their past experiences. A fundamental requirement that facilitate learning is an appropriate environment where learners can have experiences of intuition decision. This emphasizes the role that appropriate pattern of intuition and experiences play in the learning process. An important style of knowledge acquisition reflects the model to the characteristics of the heuristics learning behaviour (He and He, 2008). Heuristic refers to any techniques that improves the average-case performance on a learning task but does not necessarily improve the worst-case performance (Tao and He, 2009). It uses knowledge of previously tried solutions to guide the search into fruitful areas of the search space. The conclusion from this is that heuristics uses estimations based on domain knowledge which are represented in the form of patterns, networks, trees or graphs. Simon and

Frantz (2003) explains the approach for handling intuition in the form of the novice user and the expert user. He described intuition as subconscious pattern recognition and adds that knowledge and past experience are very important for intuition to be accurate. The drawback of this work is that he did not explain the concept in terms of how it mapped to the problems and the evolving nature of entities as well as that of the environment. Therefore, lacks practical implementation. Dundas and Chik (2011) present an implementation of human-like intuition that used series-based model and the principles of connectivity and unknown entities. In this approach, they represented the problems and experience as sets. In particular, the space of intuition contains relational mappings between experience set elements and their associated attributes. A mapping function associating elements from the knowledge set with the experience set is at the core of any problem solving activity based on artificial intuition. In general, the algorithm goes through the following stages:

1. Identification of initial conditions
2. Assessment of suitable equations from recollected pieces of human experiences
3. Analysis of relevant weight values
4. Recognition of redundant information
5. Application of adjustment factors on all considered processes to calculate the final answer
6. Check if there are any external influences and then present the output to the user.

In Anderson (2007), the author describes a sub symbolic mechanisms programmed as algorithms of low computational cost to solve elemental problems in a bizarre world. Anderson (2007) argued that for machines to achieve what humans possess, the path must include artificial intuition.

In Srdanov et al. (2016), the authors presents an artificial intuition method that searches for optimal path by combining trial and error with a random choice. More specifically, they combined logic and randomness.

The approach first forms a large set with multiple repetition of some elements which contains possible candidates for the solution. It then searches by using logic and applying random choice to select from many of its elements, which enables the solution to be reached. They argued that using random choice reduced the search space and enable the system to converge to the solution in fewer steps than without the random choice. They applied the algorithm on the case of finding a pre-conceived five-digit numbers (quintuplet) from a relatively large set of possible quintuplets. The goal is to find the correct sequence of the opponent’s quintuplet. Similarly, Srdanov et al. (2017) described the same trial and error with randomness algorithm as in Srdanov et al. (2016) and in addition used symmetry properties in pattern. They applied this to a case where the digits in a quintuplet can be repeated as against Srdanov et al. (2016) where digits in the quintuplets cannot be repeated. The fundamental difference is that with repetitions of digits, the number of moves necessary to find initial quintuplet is at most 8, while without repetition, it is possible to find the unknown quintuplet in a maximum of six attempts.

In Diaz-Hernandez and Gonzalez-Villela (2017b), the authors propose that human intuition is embodied in three stages: Inputs, Processing, and Outputs and correlated each stage to its equivalent in artificial intuition. They presented an artificial intuition approach that focused on synthetizing algorithms that improve robot performance in pick and place tasks. The approach focused on simplifying the processing stage of decision-making process, by reducing the complexity of the set of instructions needed to solve the task. The algorithm presented by Diaz-Hernandez and Gonzalez-Villela (2017b) is embodied as mathematical expressions that directly model obstacle avoidance in robots. The algorithm follows: *a)* Collect data from human abilities achieved by intuition, *b)* Analyze implicit characteristics of the intuitive performance (such as patterns, templates, among others), and *c)* Synthetize the algorithm that emulates the intuitive action/reaction. They showed an experiment that reduced the execution time of the task aided by the algorithm and also augmented its accuracy.

Tao and He (2009), suggested a learning system based on intuition

through artificial intuition networks as a mechanism of cooperative learning which will provide to the user a game-like experience, making situations more obvious and easier to learn. They developed a general instrument for measuring trusted intuition in the context of intuitive learning system. The drawback of the system is that they only described principles but did not describe an implementation.

Liu and He (2019) provided a model for constructing inference machine of criminal investigation. Their model is based on the principles of memory mapping and perceptual inversion. They used an intuitive characteristic index to establish humans intuitive reasoning. They defined experience and used the model to analyse a crime investigation case and showed a reliability result of 85% in practical application.

A related research area is computational creativity. It involves exploring computational approaches to simulate and evaluate creativity, using AI techniques such as semantic networks (Boden et al., 2004). Zhang and Zhong (2016) stated that creativity blends seemingly disparate ideas often embodied in existing knowledge. Boden et al. (2004) hinted that the creative mind searches through a search space. According to Gilovich et al. (2002), the search space can be traversed during creativity by using different thinking styles strategy defined by rules and constraints that makes analogical pattern matching possible. This include unconscious, associative heuristics such as intuition, rules of thumb, trial and error and common sense. They stated that a conscious thinking style can be constructed from a logical rule such as find as many as possible or find the most unusual, combined with a constraint. An important aspect of creative models is the use of conceptual blending (Koestler, 1964) that explores how concepts can be imagined and combined into new meaning. Such combination can be based on similarity measures. Blancke and De Smedt (2013) presents a model that exhibit creativity in an artistic context, that are capable of generating or evaluating an artwork. They implemented a concept search space as a semantic network of related concepts, and search heuristics to traverse the network in order to generate creative concepts. Pease and Corneli (2018) applied a combinational meta search approach that uses search over the space of all

possible models for the class of artifact desired to implement computational creativity. They used the technique of concept blending, amalgamation, and compositional adaption to allow the recombination of items from within a knowledge base, while retaining some of the structure from the parent items. Kelly et al. (2008) showed an interactive genetic algorithms (IGAs) approach that simulated creativity through the use of divergent and convergent thinking processes. In the approach, the convergent process hones in on specific designs, while the divergent process explores design possibilities in a fashion beyond pure mutation techniques typically used to introduce population diversity. It uses the Monte Carlo simulation to explore the effect of merging two populations developed by the divergent and convergent methods.

### *2.3. Overview of the application of artificial intuition*

In this section the author present an overview of some of the potential applications of model of Intuition. The author have categorized the overview of the applications into *a)* Integration with Clinical Decision Support (CDS), *b)* Simulating the decision of a dealer in the business of buying and selling of stocks, *c)* Simulating the decision of a player in a game.

#### *2.3.1. Intuition and Clinical Decision Making (CDM)*

Decision making in a health care context is referred to as clinical decision-making and it is an integral element of the process of patient management. The application of intuition in clinical decision making in practice has been acknowledged.

Intuition in clinical practice is the process that consists of rapid establishment of associations between information collected from the patient and pattern recognition that becomes so automatic as to be unconscious (Billay et al., 2007; Buckingham and Adams, 2000). Lyneham et al. (2008) found that the intuitive decision-making process of clinical professionals runs through three phases: cognitive, transitional and embodied intuition. The cognitive intuition results from the combination of knowledge and experience among clinicians which can occur on both conscious and unconscious levels. Clinicians use this cognitive connection to perform clinical activities.

The transitional intuition is the consequence of cognitive and embodied intuition where clinicians make a logical transition of these two aspects. The embodied intuition is the result of utter trust of clinical practitioners on experience and knowledge that build confidence to perform interventions in critical situations.

There are increasing evidences from research that the intuition in nursing has played an important role in clinical decisions and thus ensuring patient safety (Aflague and Ferszt, 2010). Tanner (2006) present a study where he conceptualizes clinical reasoning as the process by which clinical practitioners make clinical judgments by selecting from alternatives after weighing the evidences using intuition and by pattern recognition. Within the clinical context, Langan-Fox and Shirley (2003) found that practitioners make intuitive clinical decisions in situations with high levels of uncertainty, where there is no previous precedent, when facts are limited or not pointing toward a particular course of action, and when there is limited time and pressure to arrive at a decision. Similar result was expressed by Thompson et al. (2004) and he added that using intuitive reasoning is particularly useful when the number of judgment steps required for selecting an intervention for patients with chronic conditions lacks complete information and treatment needed to be delivered in a short frame of time.

In a study by Brien et al. (2011) on the use of intuition in homeopathic clinical decision making, they found that Intuition was used to enhance the practitioner-patient relationship and for prescribing decision.

### *2.3.2. Intuition and Business Decision*

There are compelling evidences from research that intuition finds significant application in strategic business decision making (Miller and Ireland, 2005). Simon (1987) presents a descriptive research on managerial decision making and problem solving where he provides insights into the nature of intuition. He noted that executives who attend to real-time information are actually developing their intuition and aided by intuition, they can react quickly and accurately to changing stimuli in their firm or its environment. Although the data are limited, the executives who relied most heavily on

real-time information were also most frequently described as being intuitive. Simon explained that intuition is dependent on past knowledge and experience for better recall of solutions to the given problems or normal logical process. Isenberg (1991) identified a number of ways in which executives use intuition: sensing a problem; performing pre-programmed behaviour patterns; producing an integrated picture; as a check on rational analysis; and as a way to by-pass analysis. Arguably therefore, becoming a more effective decision maker is about learning to make sense of intuitive feeling, judging when to trust it and feeling confident enough through practice to use it. Mousavi and Gigerenzer (2014) suggest that humans elect to make a snap judgement based on intuition, rather than deliberating with available information and in making such judgement, the thought process create a decision tree that starts with the following fundamental question:

1. If the worst-case scenario of a proposal were to occur, could I survive?
2. If no, don't pursue it.
3. If yes, the next question might be whether the company was well-positioned as a first mover in an area.

By making each decision sequentially, the company can more effectively limit its information to relevant factors, thereby avoiding information overload and not attempting to quantify the unquantifiable. Miller and Ireland (2005) presented a study that draws evidence from the fields of behavioral decision making, strategic decision making, and mental modeling and used that to evaluate intuition's costs and benefits in light of an organization's goals. They found that executives use intuition in their decision making. However, they noted that the use of intuition can sometimes be troublesome decision tool. They offered tactics that decision makers can use to make intuitive judgments and choices less troublesome.

In Hensman and Sadler-Smith (2011) study, they found that executives' reliance on intuition is not only related to time and uncertainty but also organizational contextual factors. This means contextual knowledge is important in making an intuitive decision. They provided a conceptual framework, and a typology of intuitive and contextual signalling which provide



bases for practical recommendations for managerial cognition and decision making behaviour.

Lipshitz and Shulimovitz (2007) present a study of how loan officers' intuitions affect their credit decisions. They found that the loan officers analysed credit applications as both a financial transaction and an interpersonal interaction. As a consequence, they integrate hard financial data regarding the soundness of the application with soft impressions and intuition regarding the credibility of the loan applicant. They noted that the intuition can be traced to observable specifiable cues extracted from the clients' behaviour. These cues are the significant patterns that feeds the intuitive decision.

Similarly, La Pira (2011) validated the importance of intuitive decision making to successful entrepreneurs. They found that entrepreneurs identify opportunities based on cues from the environment that they filter and process through a number of mechanisms. but they however, noted that entrepreneurs also make use of rational decision-making in addition to intuition.

This study considered a use case of the intuitive feeling decision making process of a successful entrepreneur in the stock/shares market domain. The process considers the knowledge from past experiences and the contextual information. There are two possible options in the decision. It is either to proceed or not to proceed. We define a relationship between the choice and action. Then the relationship is enriched with information from past knowledge. The essential decision options in the stock market domain are: Buy; Not Buy, Sell, Not Sell. See table 1

S/N	Decision Options	Evaluation Value
1	Buy	1
2	Not Buy	0
3	Sell	1
4	Not Sell	0

Table 1: Decision Options and Values

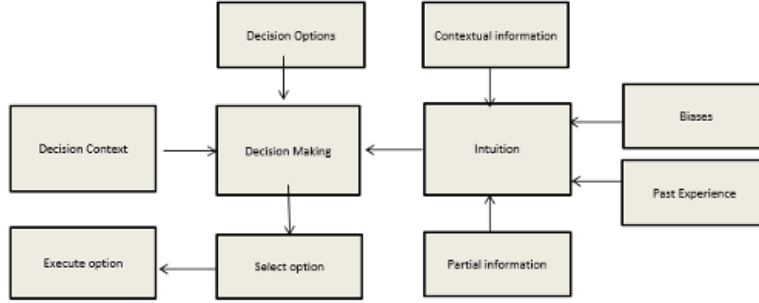


Figure 2: Overview of entrepreneur decision making process

Conceptually, the entrepreneur decision making approach consists of a number of steps. In general, within a given decision context, every decision option gets an evaluation value associated with the intuitive feelings for the same decision context and option. This intuitive feeling is informed by past experience, contextual information and enriched partial information. The essential decision options are: Buy; Not Buy, Sell, Not Sell. Figure 2, gives a high-level description of the various steps involved in the decision process of an entrepreneur. The central process is the decision-making process. The decision context kick starts the entire process. The decision context captures the situations that a broker wants to make a decision about given the multiple uncertainties and volatilities in the market. The decision-making process gets inputs from the available options and choices to pick from, and the intuitive feeling. Furthermore, the intuitive feeling get inputs from past experience, contextual information and enriched partial information. Finally, the output of the decision process which is the selected decision option is then executed. Subsequently, the outcome of the decision is evaluated against the considerations for the success of the decision process. The fundamental considerations for the success of the decision process are:

- Feeling good about the decision
- Positive state of mind
- Successful outcome of the decision

### *2.3.3. Intuition and Player in a Game*

An important application of artificial intuition is in the area of games. The ability to remember past events in the form of patterns and to recall and superimpose that pattern on present situation during decision making is an important feature of artificial intuition. For example, after winning or losing a game, one tends to remember the last moves. These moves are captured as a mental model. In similar game situations, this mental model, captured as a semantic network is assessed to get the relatedness and similarity with the current situation. More specifically, there is a mapping of the initial knowledge in solving similar problem in the past and if the overall knowledge of the current scenario is consistent with the initial assessment, then it can be assumed that it is accurate and no further analysis is suggested. This suggests that the intuitive-feeling associated with the prior knowledge is an appropriate modelling representation of the corresponding scenario. Consequently, the player no longer plays the losing strategy in similar situations. Dundas and Chik (2011) showed an example that used intuition to predict the possible hand combination in poker hand game.

Silver et al. (2016) showed how they captured human moves and replicate the intuitive pattern recognition in AlphaGo. This intuition is used in judging how good a particular board position is during game competition. Liu and He (2019) provided the result of an artificial intuition model that was used to analyze a crime investigation case. The model was based on the principles of memory mapping and perceptual inversion.

### *2.4. Emotion and Decision Making Approaches*

Changing emotional states is an important aspect of human behaviour. Studies in cognitive science, neuroscience, and psychology have extensively investigated phenomena relating to the effects of emotion on the mental process of acquiring knowledge and the understanding through thoughts, experiences and senses. Moreover, the computational exploration of the interactions between emotion and other cognitive processes is important for developing architectures for general intelligence, and in particular for decision systems. Some studies have cast the computational modeling of human

decision making as an interdependent set of deliberative and associative cognitive processes (Madni et al., 2006). The associative subsystem incorporates emotional heuristics and is described as fast, intuitive, associative, and difficult to control or modify, while the deliberative subsystem incorporates slower, rule-based, serial cognitive processes (Carruthers, 2009). Both systems are an advancement on the two-system cognitive process theory of human reasoning proposed by Kahneman (2002) as system 1 and system 2. Moreover, system 1 may be likely susceptible to the effects of emotional state, which may get in the way during execution of the decision (Gasper, 2003; Dreisbach, 2006). In other words, emotion can potentially refocus cognition away from a task at hand, thereby causing distraction, even though the emotional episode is irrelevant to the task at hand, nonetheless, system 1 is very useful during critical and time based decision making scenarios (Slovic et al., 2007). This study subscribes to this view.

In the remaining part of this section, the author presents emotion and decision making approaches. In the emotion and decision making space, a variety of approaches have been proposed, which will be discussed in this section. From the artificial emotions perspective, Minsky (2007) have explained that emotions are an integral part of human cognition of the everyday world. Minsky has gone further to suggest that much of people’s affective attitudes and knowledge is an integral part of their commonsense model of the world. Minsky regarded minds as the collections of vast number of semi-autonomous, intricately connected agents, regarded as the society of agents. He explained how minds relate to such functions as motivation, language, memory, learning, intentions, and metaphors. He modeled the function of mind from the view of computer architecture. On the contrary, Dreyfus (1972) have argued that human expertise depends on unconscious instinct rather than conscious symbol manipulation and on having a feel for the situation rather than explicit symbolic knowledge. Hart (2009) regards this as the interplay between the conscious mind and the intuitive mind.

Rauterberg (2010) argue for emotions as the main characteristic of the communication. He considered emotions as the conscious experiences of affect based on complex internal states. He distinguished between intuition

and emotion. While he considered intuition as the powerful information processing function of the unconscious, he noted that emotion is the result of this process communicated to the conscious. Emotions are the perception of the mapping from the high dimensional problem solving space of the unconscious to the low dimensional space of the conscious. Rauterberg (2010) showed an example of a different design and architecture for entertainment robots and other entertainment products with ‘emotional’ behavior

Velásquez (1998) presents a computational emotion-based control for autonomous robots, where specific mechanisms of emotions as building blocks are applied to acquisition of emotional memories that serve as biasing signals during the decision making process and selecting actions. Human emotions are mental states of feelings that arise spontaneously rather than through conscious effort and are accompanied by physiological changes in facial muscles which implies expressions on face. Some of critical emotions are happy, sad, anger, disgust, fear, surprise etc. Facial expressions play a key role in nonverbal communication which appears due to internal feelings of a person that reflects on the faces. Velásquez (1998) approach models six different emotions for decision-making: anger, fear, distress/sadness, joy/happiness, disgust and surprise. It considers natural (or innate) releasers and he also included the capacity of acquiring learned releasers. The natural releasers are, for example in the case of fear, situations in which its sensory systems would not work properly (dark environments) and the detection of archetypal predators. The learnt releasers correspond to the stimuli that they tend to be associated with and predictive of the occurrence of natural releasers. Velásquez (1998) shows a scenario of an emotional memory that is formed when a person interacts with a robot. In this scenario, petting and disciplining of a robot are natural releasers for the joy and fear emotional systems, respectively. So, when the fear emotional system becomes active because of the punishing action, the present stimuli from the presence of a person is determined and this is associated with the fear emotional system as a new learned releaser. Subsequently, this learned releaser acts as biasing mechanisms during the action-selection process when similar situation is encountered by the agent in the future. Velásquez (1998) approach is

based upon previous work of Damasio (1994) that shows that intuition and emotions play crucial roles in the ability to make smart, rational decisions. Basically, humans acquire emotions over time, and these emotions serve as a signal that influences decision making process.

Gmytrasiewicz and Lisetti (2001) use decision theory to model choice and decision biases. They modelled processes such as state utility shifts, state evocation probability shifts and allotted planning time. They incorporate how emotional state changes affects action tendencies. For example, they show how negative emotional state could result in negative outcomes.

Similarly, Hudlicka (2007) show a model of attention focus shift, decision biases and cognitive ability based on current emotional state. The approach called MAMID, modelled a system-wide working memory capacity and inference speed parameters altered. They also modelled read/write emotional parameters on mental constructs

Also, Marinier III et al. (2009) presents a model called Soar-Emote where they modelled planning goal shift and abandonment in addition to attention focus shift. The approach modelled the process of reinforcement learning and recall biases as well as emotional states.

Ahn (2010) of MIT used modelling processes of experienced-utility and expected-utility function parameters change and show how different emotions affect risk-aversion when choosing actions. Ahn (2010) shows that negative emotions made people more risk-averse in the face of gains. When it comes to losses, anger made people more risk-averse and fear more risk-seeking. in addition, they show how an agent’s previous emotional experience can be leveraged upon for predictive purposes.

Sloman (2001) presents H-CogAff, an architecture that model emotional state-based decision biases. The architecture incorporates the potential for pervasive emotional effects, as well as a deliberative layer of cognition that includes decision. The system has an oversight mechanism for sensing pattern-driven “alarms” from all levels of its cognitive processing: reactive, deliberative, and reflective.

Murphy et al. (2002) propose emotion-based control for multi-agent systems. The system provides the ongoing monitoring function and from that

monitoring, emotional states are generated. The implemented artificial emotions such as happy, confident, concerned, and frustrated and each of these emotions is released depending on the task progress.

Similarly, Salichs Sánchez-Caballero and Malfaz Vázquez (2012) show an approach where autonomous agent makes decisions based on drives, motivations, and emotions. In this approach, the agent has certain drives within a certain range and the motivations are understood as what moves the agent to satisfy a drive which is a particular emotion. Basically, agents have motivations that drive it to satisfy certain emotions that help in its wellbeing. They modelled and evaluated the emotions of happiness, sadness, and fear as the wellbeing of the agents. Thus, emotions indirectly drive behaviours that lead individuals to act, achieve goals and satisfy needs (Sequeira et al., 2016).

Lerner et al. (2015) propose an integrated model of decision making that implements rational thinking and emotional inputs. They categorized how emotions influence decision making into 8 themes: Integral emotions, Incidental emotions, emotions via the content of thought, emotions via goal activation, emotions in depth of thought, emotions in interpersonal decision making. They argued that unwanted effects of emotions on decision making can sometimes be reduced. Integral emotions are the emotions that arise from the judgment or choice at hand. Lerner et al. (2015) stated that integral emotions can act as a beneficial guide or a bias towards decision making. They identified anger, sadness, fear and disgust as forms of negative affect while pride and surprise as forms of positive emotions. They further argued that fearful people tend to see greater risk and angry people tends to see less risk. Moreover, emotions potentially bias the value attached to outcomes. For example, intense negative emotions enhance valuation of short-term outcomes regardless of negative long term consequences. This however depends on the particular domain of judgement or decision making. Lerner et al. (2015) and Fenton-O’Creevy et al. (2011) hinted on some ways to minimize the harmful effect of emotional response on decision making, such as avoid situation, time delay, suppression, reappraisal or modification of situation. Time delay strategy is based on the assumption that emotions are short-lived

as physiological responses quickly fade.

De Melo et al. (2012) presents a computational model that captures the social function of emotion between an agent and its surroundings and what human infer based on the emotions displayed, how the agent is appraising the ongoing interaction with the environment. They learnt the model's parameters by using empirical data from an experiment where people engaged in a social dilemma with embodied agents that expressed emotions. They used a virtual character that was able to portray emotions when playing the iterated prisoner's dilemma game. The virtual character's facial expressions were used to enhance human decision prediction. De Melo et al. (2012) discussed some agreements that are critical for emotions to be considered: joy occurs when the event is conducive to one's goals; anger occurs when the event is not conducive to one's goals, and is caused by another agent and one has power/control over it; sadness occurs when the event is not conducive to one's goals; guilt occurs when the event is not conducive to one's goals and is caused by the self.

Fenton-O'Creevy et al. (2011) show how emotions influence perception and performance of traders in investment banks. They found two types of emotion regulations strategies: antecedent-focused and response-focused and conclude that traders with antecedent-focused emotional regulation strategies achieve a performance advantage over those employing primarily response-focused strategies. This however depends on the level of emotional experience of the trader.

Categorization of emotion induced decision making Based on the review, this study have extracted and categorized emotion based decision into 7.

- Decision making based on "safety concerns" induced emotions (Luce, 1998).
- Decision making based on "frustrated anger" induced emotions (Leith and Baumeister, 1996).
- Decision making based on "fear" induced emotions (Lerner and Keltner, 2000; Salichs and Malfaz, 2012).



- Decision making based on "sadness" induced emotions (Lerner and Loewenstein, 2004; Salichs and Malfaz, 2012).
- Decision making based on "normal processing" induced emotions (Bechara et al, 1999).
- Decision making based on "happiness" induced emotions (Salichs and Malfaz, 2012).
- Decision making based on "disgust and surprise" induced emotions (Velasquez 1998).

With such categorization, researchers have been able to extract and determine instances of possible emotions and how it affect decision making. For example, Gratch and Marsella (2005), used negative emotional state such as displeasure to provides a straightforward correlation to narrow-minded but careful decision making, while positive emotional state such as pleasure have been mapped to broad decisions that attempt to achieve multiple goals with less attention to detail and more heuristic processing. When a person makes a wrong or right decision, it is possible that the person is in 'normal' particular frame of mind. The assumption can be made that the person is biased and therefore is in a good state of mind.

Detail discussion of emotion is outside the scope of this work as emotion is a complex subject that crosses several fields of studies which includes Psychology, Neuroscience, Artificial Intelligence and Philosophy. In each of these fields, one will most certainly find many perspectives on the study of emotion and different schools of thought, whose ideas are sometimes opposing. This study takes the computer science perspective and emphasis on computational modeling of artificial intuition.

The above reviews show that there is evidence to suggest that emotions play an important role in decision making. Most of the models and frameworks of emotions are specifically on modeling of categorical emotional states and cognitive science (Velásquez, 1998; Sloman, 2001; Murphy et al., 2002; Naqvi et al., 2006; Salichs Sánchez-Caballero and Malfaz Vázquez, 2012). More specifically, some of these studies use emotional state to provide direct

quantitative correlation with emotional effects on decision-making (Murphy et al., 2002; Salichs Sánchez-Caballero and Malfaz Vázquez, 2012; Lerner et al., 2015; De Melo et al., 2012), while others use emotion to influence cognitive processes by means of providing goal-based cues and biases (Velásquez, 1998; Sloman, 2001).

Although, this study is not about modelling categorical emotion and decision making, the author have presented this section to provide the evidences to show that human processes are susceptible to the effects of emotional state. This is part of a wider line of inquiry into artificial intuition and decision making. The justification for this is that there is a resonance between the emotional content of memory and the overall emotional state and this act as a cue to the intuitive feeling. This is part of the associative-semantic network. More specifically, the emotional state of a person is one of the components that feeds into the intuitive decision making process. The author do not intend to pursue this discussion any further. It was presented out of theoretical interest and curiosity and to show some evidence.

### *2.5. Rational Decision-Making Approaches*

In the first part of this section, the author presents the manual and automatic creation of Bayesian network. The second part will give a brief discussion on deep learning and then present rational thinking based decision making approaches that basically use deep learning at the core of the methods.

Bayesian Networks (BNs) are directed acyclic graph that represents joint probability distribution over a set of random variables (Ben-Gal, 2008). More specifically, it is a graphical model (structure) that is used to represent knowledge about an unpredictable, imprecise and uncertain domain (e.g decision situation domain). Here the nodes in the graphical structure represent the random variables in the domain while the edges between the nodes represent conditional probabilistic dependencies among the corresponding random variables.

There are basically two levels at which BN is modelled: qualitative and quantitative. The qualitative level describes a directed acyclic graph in

which nodes represent variables in the uncertain domain, and directed arcs describe the conditional independence relations embedded in the model. The quantitative level describes the dependence relations in terms of conditional probability distributions for each variable in the network. The probability distribution for each variable is captured in a Conditional Probability Table (CPT).

The CPT describes the likelihood of any node in the Bayesian Network being in one state or another without current evidence and in particular they depend on the causality relationships between some nodes often described by prior information on such networks. There is one CPT for each node, which describes the conditional probability of that node given the different values of its parents (Friedman and Goldszmidt, 1998). The complete joint probability distribution (JPB) for the network is expressed by the CPTs for all the variables together with the conditional independences described by the network. In other words, using the CPT for each node, the joint probability distribution of the entire network can be derived by multiplying the conditional probability of each node.

#### *2.5.1. Bayesian Network Creation from data*

This section discuss how to create a Bayesian Network i.e given a training data set and prior information (e.g expert knowledge, casual relationships), estimate the graph topology (network structure) and the parameters of the JPD in the BN. In general, the steps involve in creating BN includes the following

- Identify the nodes (variables) of the domain. This task is influenced by whether the construction of the BN uses data-based approach or knowledge-based approach
- Identify relationships (probabilistic dependencies among the corresponding random variables). This task as well as the first task requires expert knowledge or significant linguistic or world knowledge.
- Identify the values of the nodes. Nodes could be discrete or continuous values. For discrete nodes, typical values are Boolean (True or False),

ordered (flow, medium, high) or integral values (1 to 120). Boolean nodes, typically represent propositions, taking the binary values true (T) and false (F). The values for each node should be both mutually exclusive and exhaustive, which means that the variable must take on exactly one of these values at a time.

- Model the BN structure that captures the joint probability distribution over the identified set of random variables.
- Identify the conditional probability table (CPT) for each child node for different combinations of the states of its parents. Discrete variables normally take the form of a CPT.
- Apply reasoning to the BN to draw inference by analyzing the probabilities of each variable in the network

BN can be created manually or automatically. In the manual method, a researcher uses domain knowledge and experience and applies both structured and unstructured techniques to elicit domain information and subsequently learns the network structure. More specifically, the manual method is based on making statistical inference tests to detect the independence relationships in data that are represented by the graph. Amirkhani et al. (2016) exploits the opinions of multiple domain experts regarding cause-effect relationships to learn BN. Similarly, Nadkarni and Shenoy (2004) describe a systematic procedure to construct Bayesian networks from domain knowledge of experts using the causal mapping approach. They applied this in the context of an information technology application outsourcing decision. Fenz et al. (2009) exploits the ontological knowledge base to model threat probability determination in the security domain. Automatic learning of Bayesian network from data involves constructing a network model that best represents an underlying data structure based on some algorithms. Cheng et al. (1998) classifies Bayesian network learning approaches into:

- Search and scoring methods
- Dependency analysis methods.

Search and scoring methods consist of choosing among the possible network structures within a given number of nodes and then estimate the probability distribution associated with the chosen structure. Such networks are then used to predict future events. These methods use various metrics such as maximum likelihood estimates to optimize the likelihood of a good topology. A drawback associated with the Search and scoring is that the problem of finding the graph optimizing the score is NP-hard with a search space, which is the set of all directed acyclic graphs (DAG) of a given size. This search is super exponential, in particular when a large number of parameters need to be searched in a distributed network where there are multiple nodes. This would incur huge amount of communication overhead to synchronize nodes when these huge parameters are searched are updated (Ooi et al., 2015; Oviedo et al., 2016). This motivates the need to find fast and optimal approaches because of the super exponential challenges. The idea of dependency analysis methods is to compute a dependency measure for pairs of nodes in a dataset and use a closeness measure, for example, to determine whether an edge could be placed between the pairs of nodes. A drawback of this approach is that since independencies are only partial properties of the data set, the dependency approach usually cannot learn the full structure (Xu and Srihari, 2014). Example of dependency analysis method was presented by Cheng et al. (1998). Cheng et al. (1998) applied information theoretic measures to identify mutual closeness of pairs of nodes. Some of these algorithms used to implement BN learning requires complete ordering of the variables and some require partial ordering and some do not. In most cases the result is either to determine conditional dependencies or independencies.

In summary, the manual methods are based on domain experience to learn the network structure, while the automatic method is based on current data, learning structure from calculation automatically as well as some prior knowledge. However, in order to have accurate system, such prior knowledge usually requires extensive pre-processing and annotation. A typical automatic method is the Search and Score based algorithms that takes a dataset  $D$  and priori knowledge and then performs a Bayesian scoring,

greedy hill-climbing search and returns a BN structure with the highest post probability. Since this is heavily based on data, if there are missing data in the dataset, these set of data have to be imported, imputed or estimated from other sources (Lucas et al., 2004). Also, there has to be enough data to satisfy the algorithm’s requirements for reliable estimates of the conditional probability distributions. For the manual construction, the conditional probability distributions are assumed to be a priori knowledge. Automatic learning involves both the network structure creation and conditional probability distributions estimation.

The study of deep learning has a historical perspective (Schmidhuber, 2015) and many deep learning studies have been done and applied in diverse areas of information processing such as speech recognition, speech synthesis, and audio processing (Deng and Yu, 2014; Schmidhuber, 2015). For example, Torralba et al, (2008) presented a deep network approach for performing image recognition. Hörster and Lienhart (2008) explored deep networks for deriving a low-dimensional image representation appropriate for image retrieval. They used network consisting of multiple layers of features to capture higher order correlations between basic image features. They evaluated the approach using real world large-scale image database and compare it to image representations based on topic models. Ooi et al, (2015), presented an intuitive model based on the layer abstraction for a distributed deep learning system. The system called SINGA, supports both synchronous and asynchronous training frameworks and trains big models over large datasets. Over the last decades, deep learning has become the focus of significant research in machine learning (Bengio, 2009), which has demonstrated to have a wide impact on many multidisciplinary fields and research areas (Arel et al, 2010; Bengio et al, 2013). Loosely speaking, deep learning sits between numerous research topics, including neural networks, artificial intelligence, optimization, and pattern recognition, decision making to name but a few. Several similar, yet slightly different definitions of deep learning have been presented. In this study, the author refer to deep learning as a set of models with multiple layers of information processing, based on supervised or unsupervised learning of feature representation, where all the parameters

are fully investigated to provide an exhaustive modelling approach (Bengio, 2009). In a deep learning approach for learning a BN, the emphasis is to follow a more precise and logical approach, where as many nodes as possible are checked in order to learn the network. Here, the depth of credit assignment paths, which are chains of possibly learnable, causal links between actions and effects are considered to be deep and therefore, exhaustively modelled (Schmidhuber, 2015). More specifically and in the context of decision making, as much nodes as possible are checked nonintuitively, in order to make a decision on a large network. The whole process is normally nonintuitive.

Complex networks with different uncertainty layers have been extensively used to describe and model a variety of scenarios within the decision-making domain (Trovati, 2015). Examples of such networks include Dependency Networks and Influence Diagrams, which are characterized by the mutual relationships between nodes representing specific concepts. However, the identification of such concepts and their relationships is usually a complex task, as a variety of probabilistic and topological constraints need to be addressed, especially when extracted from textual sources (Blanco et al, 2008). Furthermore, depending on the scenario, there are potentially a large number of parameters, which increase the overall computational complexity (Trovati, 2015). As a consequence, when deep learning has been applied to this context, the extraction of actionable information from such a scenario may raise a variety of issues (Trovati, 2015; Trovati et al, 2016).

In the remaining part of this section, the author present rational decision making approaches that applied deep learning at the core of the methods.

Jameson et al (2000) present an approach where an adaptive intelligent interface makes a decision on action to perform in a given situation based on a derived Bayesian model from empirically validated data. The approach learned the influence diagram from the derived BN and used that for making the adaptation decision automatically. One of the main issues with this approach is that it is usually impossible to obtain experimental data in a situation that is identical to the situations in which the system is actually to be used. Another issue is that by bundling instruction to make adaptation decision, the working memory may be overloaded thereby causing

performance issues especially for a large network.

Ivanovska and Giese (2011) proposed a decision making approach based on a formal logic. The main idea here is that they treated the choice of an optimal decision as a problem of logical deduction. They described syntaxes and semantics for the application of logic to decision making. In their approach, they let each complete sequence of actions impose a separate probability measure on a common set of worlds equipped with a utility function. The formulae of the logic may refer to only a subset of the decisions, which allows for a more compact representation in the presence of independencies. They argued that their formalism can deduce at least the same consequences as are possible with the most popular graphical approach. However, they did not describe any problem case study.

Bhattacharjya and Shachter (2010b) transforms influence diagrams into decision circuits and use the decision circuits to compute the optimal decision policies. The method assumes that all uncertain variables in the model are represented by discrete probability mass functions (pmfs) and that decision variables are discrete. Their method exploits the local structures present in an influence diagram, such as deterministic relations. The drawback of this approach is that it starts by solving small low-level decision problems and gradually build on the results to solve larger problems until the solution to the global-level decision problem is found. This can potentially waste computation time in solving decision scenarios that have zero probabilities or are unreachable from any initial state by following an optimal decision policy (Yuan and Wu 2010). Also, the larger number of parameters to access at the global-level, the more communication overhead it will incur in order to synchronize the low-level and global-level states.

Kant and Thiriot (2006) propose an agent-based model of human decision making. In the model called CODAGE (Cognitive Decision AGEnt), the decision maker is modelled by an entire multi-agent system, where each agent is in charge of a particular sub-process of the whole decision. The main idea they present is to distinguish between relevant and irrelevant information coming in from the sensor layer by using experiences from the past to determine the priorities. Here they described an approach where



the decision-making process comprised of three phases: intelligence, design and choice. Intelligence is the process of understanding what the decision is about by exploring the current context in respect to the constraints of the decision. In the design phase, the possible different alternatives in the form of solutions to the problem are created and finally in the choice-phase the system chooses one of these solutions or alternatives using heuristics. The approach was implemented by setting up a tree of alternatives, where the nodes represent different possible states of the world in past, present and future. The arcs between the nodes represent possible transitions of these states. This tree is subsequently modified and enhanced by different agents. The main drawback is that it is very computationally intensive, and it is non-trivial to synchronize the jobs of the agents in the sub-processes. Also, it is more useful for studying a single decision rather than a real-time decision process.

Bencomo et al (2013) presented an approach where they combined Bayesian Network and decision network to support the decision-making of self-adaptive systems. They describe the mechanisms to cope with uncertainty and automatically make the best decision in self-adaptive environments. This approach is largely a deep BN and it is computationally intensive to calculate the expected utilities associated with a decision and the evidences in a large network. Detwarasiti and Shachter (2005) presented an influence diagram approach to model team decision making under uncertainty. The approach assumed that all team members agree on common beliefs and preferences, but complete sharing of information is generally impossible. Therefore, they represented the team as a single rational individual with imperfect recall and then used the solution concept of  $k$ -stability and the corresponding solution method of strategy improvement to evaluate the team decision situation having incomplete sharing of information. In general, their approach is a  $k$ -neighbor local search in the space of strategies and an exhaustive  $k$ -local search can only be applied to networks with less than  $n=100$  decision variables and with very small values of  $k$ . A major drawback for the approach is that it requires considerably more computational effort in the search for optimal strategy and finding the optimal strategy is NP-hard (Mauá and

Cozman 2014).

In conclusion, the ability of the rational decision making approaches to perform their evaluation based on large quantities of parameters, which define a specific scenario is at the core of deep learning.

## *2.6. Guts-feeling and Decision Making*

A related research area to artificial intuition is gut-feeling. This section presents a discussion of gut-feeling in the decision making process. The concept that the gut and the brain are closely connected, and that this interaction plays an important part in certain feeling states and in intuitive decision making (Mayer, 2011) is an important endeavour to pursue in this study. This connection between the gut and the brain is what sometimes drives the decision making process. This section is discussed in order to provide a wider line of inquiry into the computational model of artificial intuition and decision making.

Gut feeling is a closely related concept to artificial intuition and emotion. While the terms intuition and gut-feelings are often used interchangeably, the author make the distinction in this study. Intuition is an understanding or knowing of a situation without specific data or evidence at the time. In other words, analytic reasoning is not part of the intuitive process. While gut-feeling is an innate, hardwired tendency. For example, humans have biological, hardwired guts for survival and reproduction. According to Sadler-Smith and Shefy (2004), guts feeling is the skill of focusing on those potentially important but sometimes faint signals that fuel imagination, creativity and innovation and feed corporate success in globally competitive business environments. In addition, Sadler-Smith and Shefy (2004) recommend the following questions that may help to gauge the extent to which business executives can rely upon intuition in decision making.

- Do you trust your hunches when confronted by an important decision?
- Do you feel in your body if a decision is right or wrong?
- Do you put a lot of faith in your initial feelings about people and situations?

- Do you put more emphasis on feelings than data when you make a decision?
- Do you rely on your gut-feelings when dealing with people?
- Do you trust your experience when arriving at the reasons for making a decision even if you can't explain why?
- Does your intuition often turn out to have been right all along?
- What is the reaction in your organization to decisions made on the basis that they felt right?
- Do you keep your intuitions close to your chest? If so, why?

Isenberg (1991) provides a study to show how gut-feeling pattern is used to bypass the rational analysis in the decision process. He noted that the capacity to learn how the gut-feeling works and be able to trust it can potentially improve the effectiveness of individual decision making process. Essentially, the guts feeling course of action of entrepreneurs goes through the following thought process

- Do I feel happy about this choice I want to make?
- Do I feel good about this choice?
- Does this situation give me or take my energy?
- Am I going toward an adventure or running from fear?
- Am I listening to my lessons learned from the past?
- Would I make the same choice if I had a million dollars in my pocket now?
- Do I feel respected and valued?
- Am I trying to control the situation or am I leaving room for expansion?

To answer the above questions, the guts feeling decision process accesses information about the following

- Emotional state of mind
- The biases
- Additional historical information (known knowledge from past experience) that has been built-up and is held in a mental map about the subject of interest. Decisions that are made in circumstances similar to previous experience, and whose outcome could be potentially harmful, or potentially advantageous, induce a somatic response used to mark future outcomes that are important to us, and to signal their danger or advantage (Damasio, 1994). In general, somatic marker is concerned about the outcome of the choices we make about the decision options that are presented to you. Guts feeling is created based on such experiences and this aids the decision process in an automatic manner. Thus, when guts feeling assesses a decision contexts and juxtaposed a negative somatic marker to the outcome of the decision option, it sends an automatic response to drop that option.
- Partial information about the subject of interest.

In fact, human decisions are not driven by data, rather, they are informed by data and based on several variables beyond data, including somatic markers, past experiences, judgement and context. Human emotions can sometimes interrupt signals to cognitive processes and this interrupt can potentially refocus cognitive attention onto an emotionally compelling stimulus (Ohman, 2001). It has been noted that emotional intensity grants a heightened priority to relevant concepts attended to during an emotional episode (Ohman, 2001). As stated by Ritter et al (2006), cognitive cues not central to a situation at hand would be increasingly ignored under higher levels of emotional intensity than would have otherwise be if the emotional level is low. Consequently, at critical decision making time, such emotional intensity could potentially lead to overlooking subtle but important details,

or leave the subject open to misdirection and other forms of deliberate manipulation (Slovic, 2007). During emotionally arousing moments gut-feeling can be a useful heuristic for making a decision in a timely manner, bypassing deliberative evaluation when time sensitivity is a major factor in decision (Slovic et al, 2007). Limiting the time to judge a situation's risk and benefit induces a sense of urgency and increases human reliance on the affect heuristic. Under time pressure, the inverse relationship between expected risks and benefits (higher risk implying lower benefits and vice versa) becomes more pronounced in decision-making (Finucane et al, 2000). Based on the above discussion, the author believe that the guts feeling working in concert with additional historical and partial information could lead to better and optimal decision making. This is consistent with the findings of Damasio, 1994; Loewenstein and Lerner, 2003; Zeelenberg et al, 2008 and Binali et al 2010 findings that suggest that rational thinking and emotions gives room for optimal decision making. More specifically, the additional historical and partial information are the facts that serve as cues for the guts feeling to make a potentially optimal decision. The key idea here is that there is an assumption of past experience which forms a mental model of the knowledge in the past. This mental model is represented as a state in the past using a binary value. So, if we are analysing a dataset and we find a knowledge that is consistent with what has been held in the past mental model, then the assessment for making a decision is accurate. More specifically, if we can find a state in the present knowledge assessment that is consistent with the state in the past, therefore there is an influence to make a future decision and this decision is assessed to be accurate.

In this section, the author have provided a discussion of the role of gut-feeling in the decision making process. This section is discussed in order to provide a wider line of inquiry into the computational model of artificial intuition and decision making. The idea that the interaction between the gut and brain plays an important part in certain feeling states and in intuitive decision making (Mayer, 2011) is an important endeavour to pursue in this study. This connection between the gut and the brain is what sometimes drives the decision making process. In subsequent section, an approach that

model this idea will be provided.

### *2.7. Shallow Learning Approaches of Decision Network*

In a deep learning approach, the emphasis is to follow a more detailed approach, where as much nodes and mutual relationships as possible are investigated in order to make a decision on the network, whereas on the shallow learning approach, the emphasis is to bypass and estimate few knowledges as possible in order to learn the network. The fundamental difference between shallow and deep learners are the depth of their credit assignment paths, which are chains of possibly learnable, causal links between actions and effects (Schmidhuber, 2015). More specifically, in the shallow learning approach, we ‘guess’ based on some shallow information, which is far from being precise. The driving question is ‘Do you put more emphasis on feelings than data when you make a decision?’

Moody and Darken (1989) proposed a fast learning algorithm that learn the network of locally tuned processing units. The approach uses a single internal layer of locally-tuned processing units to learn both classification tasks and real-valued function approximations. The approach uses a hybrid method that combines self-organized and supervised learning to ensure only a few units respond to any given input. This was used to achieve a shallow learning of the processing units.

Hinton et al (2006) show the use of complementary priors to derive a fast, greedy algorithm that can learn deep Bayesian Network that has multiple connections. In this approach, the top two hidden layers form an undirected associative memory and the remaining hidden layers form a directed acyclic graph that converts the representations in the associative memory into observable variables such as the pixels of an image. More specifically, the learning algorithm greedily trains one layer at a time, exploiting an unsupervised learning algorithm for each layer. They argued that this approach helps to eliminate the explaining away effects that makes it difficult to infer conditional distribution in densely connected Bayesian Network that have many hidden layers (nodes).

Nilsson and Nielsen (2014) described a declarative language called Ebba

(Embedded Baysig) that uses a shallow and arrows-based embedding for Bayesian modelling. In this approach, they used the notion of computation to be an arrow representing a conditional probability distribution. Here, they used arrow combinators to intuitively show how building blocks can be used to describe networks. Here, they used “forwards” for probabilistic computation and “backwards” for parameter estimation. it uses the shallow embedding to provides a clear semantical account and ensures that only models that support estimation can be expressed.

Peraza and Halliday (2010) proposed a fast algorithm that uses a time-lagged version of the partial correlation matrix to overcome the issue of load and time consumption in the learning of Bayesian decision Network. The approach formed the matrix by concatenating the available time series data and a time shifted version of the same data. Then, the algorithm uses series of algorithm rules to evaluate Bayesian Network patterns using the Bayesian information criterion (BIC) score for Gaussian networks. Firstly, the rule looks for partially directed three node networks where only one directed edge has been found, and finally it looks for three node networks having undirected edges. Secondly, it identifies 4-node loop paths over the entire network and also identifies the directed and undirected edges. The algorithm then performs some steps to remove the undirected edges using scoring. Finally, the rule finds the direction of the remaining undirected edges using two or more networks with the highest score.

## *2.8. Neural Network and Decision-Making Approaches*

Neural network is a system of hardware and software patterned after the operation of neurons in the human brain. More specifically, they are complex mathematical models inspired by biological neural networks. A neural network consists of an interconnected group of artificial neurons, and it processes information using a connectionist approach to computation. Neural networks consist of input and output layers, as well as a hidden layer consisting of units that transform the input into something that the output layer can use. The weight are the values assigned to each of the connection. The weight of the connection indicates the value of that connection which

is indicative of the importance.

Learning in neural network is a process by which free parameters of a Neural Networks are adapted through a process of simulation by the environment in which the network is embedded. First, the network Initialize the weights to small numbers very close to 0. Then it performs a forward pass with the production of an error signal for each output neuron. Errors are then propagated back through the system, causing the system to adjust the weights, which control the network. This process repeats over and over, and the weights are continually tweaked. During the training of a network, the same set of data is processed many times as the connection weights are refined. The learning process continues until the final output at end.

Google DeepMind (Silver et al., 2016) applied deep learning in neural networks to develop AlphaGo. AlphaGo uses neural network in which connections between layers of simulated neurons are strengthened through examples and experience. it first learned by training a deep convolutional neural network to predict human moves from a database of expert games using supervised learning (the 'SL network'). it then uses policy gradient reinforcement learning to further optimize the network by simulating games of the network playing against previous versions of itself and rewarding it for taking moves that result in wins (the 'RL network'). Furthermore, it trains a separate deep neural network to map board positions to the likelihood of player 1 winning the game (the 'value network'). Thereafter, it performs a Monte-Carlo tree search to evaluate the most promising moves suggested by the SL/RL networks according to the value network and policy rollouts. This step is used to select moves in competition. Silver et al. (2016) argued top Go players use a lot of intuition in judging how good a particular board position is and they argued that they were able use neural network to capture such intuition and replicate the intuitive pattern recognition in AlphaGo.

Kim and Han (2000) proposes a genetic algorithms (GAs) approach to feature discretization and the determination of connection weights for artificial neural networks (ANNs) to predict the stock price index. These features are the generally accepted technical indicators as the signal of future market



trends. In this approach, the GA searches for the optimal or near-optimal solutions of connection values (weights) in the learning algorithm and looks for the optimal or near optimal thresholds of feature discretization for the dimensionality reduction. This helps to reduce the complexity in feature space. Kim and Han (2000) argued that genetically evolved weights mitigate the well-known limitations of the gradient descent algorithm.

Thagard and Aubie (2008) describes an emotion neural network approach that represent the overall cognitive and somatic state of the organism. They described a model of combination of neural representation interconnected neural processes that gives rise to emotion and Intuitions. They argued that conscious experience arises when the neural representations of the brain achieve high activation as part of working memory.

Similarly, Nazir and Liljenström (2016) presents a Computational model representing the neural network of the information processing of decision-making, from perception to behavioral activity. They model the dynamics of the three neural structures (amygdala, orbitofrontal cortex (OFC), and lateral prefrontal cortex(LPFC)) as well as the interactions that are significant in the decision-making process. In the model, amygdala and OFC represent the neural correlates of secondary emotion, while the activity of OFC neural populations represents the outcome expectancy of alternatives, and the cognitive aspect of decision-making is controlled by LPFC.

Bechara and Damasio (2005) proposed a neural model for economic decision, in which emotions play a major factor in the interaction between environmental conditions and human decision processes, with these emotional systems providing valuable implicit or explicit knowledge for making fast and advantageous decisions. They described the interconnectedness and the modifications of the various neurotransmitter that happens when there is a change in the body states. They argued that the implementation of decisions under certainty or uncertainty engage different neural circuitry.

Kant (1995) presents *Categ-ART*, a neural network based approach to model the actual process of human decision making and produce decision criteria that account for categorizations made by a particular subject. The system is composed of two interconnected layers (F1 and F2), where F1

receives the afferent input features and F2 is composed of monomial cells that encode monomials. The system used a supervised learning algorithm to automatically build polynomial rules that account for a particular decision making categorizations. The approach was applied to data on choices made among bank savings schemes.

Matsuda (2006), presents an artificial neural networks model of decision making based on Analytic Hierarchy Process (AHP). The author used AHP to decompose the decision making problem into three hierarchical components: alternatives, criteria, and goal. A key component of the model is that it dealt with the situation of incomplete pairwise comparisons in the case of unknown or uncertain information when the decision is made. He further showed the validity of the model with simulation of neural network of decision-making process for school selection problem.

Murtaza and Fisher (1994) presented an ANN model approach to decision making using self-organizing multi-layered neural networks. The model enables decision making on using modularization or conventional method for building an industrial process plant based on five categories of decision attributes namely plant location, environmental and organizational factors, labour related factors, plant characteristics and project risks. The NN model was trained using cases collected from several engineering consulting and client firms.

Chen and Lin (2004) proposed a new approach for solving multiple criteria decision-making (MCDM) problems based on decision neural network (DNN). The DNN is used to capture and represent the decision maker's (DM's) preference. Then, with DNN, an optimization problem is solved to search for the most desirable solution.

In this section, the author have presented some decision-making approaches that applied neural network approaches. A Neural Network is a parallel distributed processor made up of simple processing units which have natural propensity for storing experiential knowledge and making it available for use (Haykin, 1999). Neural networks can also be viewed as weighted directed graphs in which artificial neurons are nodes and directed edges with weights and connections between neuron outputs and neuron in-

puts. Neural network resembles the brain in three aspects. First, knowledge is acquired by the Network from its environment in the form of pattern and image in vector through a learning process. Second, Interneuron connection strength is used to store acquired knowledge. Third, the neurons create parameters such as intuitions feelings from the acquired knowledge from the environment and these feelings, make us act or take decisions which is basically the output of the neural network of our brains. In Neural Network, each node performs some simple computation and each connection conveys a signal from one node to another labelled by a number which is the weight. Typically, the weight represents the strength of the interconnection between neurons inside the neural network. The weighted inputs are all summed up inside computing unit (artificial neuron). In case the weighted sum is zero, bias is added to make the output not-zero or to scale up the system response. Bias has the weight and input always equal to '1'. The sum corresponds to any numerical value ranging from 0 to infinity. In order to limit the response to arrive at desired value, the threshold value is set up. For this, the sum is passed through activation function.

The big challenge of neural network is the amount of time it takes to train networks, which can require a considerable amount of computational power, in particular for more complex tasks. The other challenge is that neural networks are like "black boxes", in which the user feeds in data and receives answers. Though, users can fine-tune the answers to achieve a high predictive accuracy rate, but they don't have access to the exact decision-making process of the network. More specifically, the reasoning behind how they reach their decisions is not readily available and cannot be readily explained (Baesens et al, 2003).

## *2.9. Summary and Discussion of the Approaches*

In this chapter, the author has presented and discussed various approaches used in artificial intuition and decision making. Moreover, a brief discussion of a related field of computational creativity, emotion and decision making as well as the role of gut-feeling in the decision making process was presented. The author have also provided a discussion on some po-

tential real-world application of artificial intuition in the industries. One potential real-world application of this model is in simulating weather prediction. For example, modeling the decision to take the umbrella when the weather is cloudy, and a football fan wants to go and see a favorite football game. Furthermore, such model can find some application in the integration with Clinical Decision Making (CDM) system to reduce diagnostic errors. Clinical practitioners use Intuition as a means of bypassing tacit and formal knowledge and its outcomes can be validated by formal means. Other potential application is in simulating the decision of a dealer in the business of buying and selling of stocks as well as the decision of a player in a game.

Furthermore, the author has provided a comprehensive review of the relevant approaches of rational decision making (Jameson et al, 2000; Ivanovska and Giese, 2011, Bhattacharjya and Shachter, 2010b; Kant and Thiriot, 2006; Bencomo et al, 2013). This study noted that the ability of these rational decision making approaches to perform their evaluation based on large quantities of parameters, which define a specific scenario is at the core of deep learning. In addition, these approaches showed that deep learning models can benefit from deeper structures and larger training datasets. Although, deep learning approaches have been shown to be effective in application involving human speech, image and visual scenes, a major drawback of deep learning method is its computational complexity. Finding the optimal network model structure where the search space is super exponential is computationally a very hard procedure (Oviedo et al, 2016; Myllymaki et al., 2002), and approximation methods are generally used instead (Uusitalo, 2007). More specifically, these deep learning approaches would normally need a deeper analysis to allow the investigation of all corresponding scenarios. However, a major challenge with these approaches during decision making is coping with a large unstructured data-sets based on partial unknown knowledge, and with multiple states. The identification of such concepts and their relationships is usually a complex task, as a variety of probabilistic and topological constraints need to be addressed, especially when extracted from textual sources (Blanco et al, 2008). Furthermore, depending on the scenario, there are potentially a large number of parameters,

which increase the overall computational complexity (Trovati, 2015). As a consequence, when deep learning has been applied to this context, the extraction of actionable information from such a scenario may raise a variety of issues (Trovati, 2015; Trovati et al, 2016). Furthermore, it is non-trivial for researchers to develop and train models with deep and complex model structures with potentially large number of parameters (Ooi et al, 2015).

Another category of decision-making models discussed above are the intelligent methods that use Neural networks to attempt to simulate the human brain by collecting and processing data for the purpose of learning (Bechara and Damasio, 2005; Chen and Lin, 2004, Kant, 1995; Nazir and Liljenström, 2016; Murtaza and Fisher, 1994). An important reason why neural network is considered in this discussion is because neural network resembles the brain in three aspects. Firstly, knowledge is acquired by the Network from its environment. Secondly, Interneuron connection strength is used to store acquired knowledge. Thirdly, the neurons create parameters such as emotions and intuitive feelings from the acquired knowledge from the environment and these intuitive feelings make us act or take decisions which is basically the output of the neural network of our brains. An aspect to consider is that by applying the connectionist model and artificial intuition, the events in the mental model can be connected and represented like neuronlike units in an artificial neural network and then spreads activation through the network in a way that activates some units and deactivates others (Thagard, 2001). At the end of the spread of activation of the elements in the mental model, the active units represent elements that are accepted, while the deactivated ones represent elements that are rejected. This approach can optimize and potentially benefit the stock investment decision and prediction of the price index. But, the big challenge of neural network is the inability to have access to the exact decision-making process of the network (Baesens et al, 2003) and the considerable amount of time it takes to train networks.

However, when a decision process is manually modelled by expert only a partial amount of information is explicitly processed as they often rely on their intuition. There is compelling evidence from cognitive research that

intuition play an important role in the process of intelligence extraction, where knowledge based on historical data, personal experience and expertise drives the decision-making process, rather than deep rational decision making (Loewenstein and Lerner, 2003). Yet, none of these deep BN models that was identified consider intuition as critical in the decision making process. This provides motivation on incorporating intuition in decision making systems. Therefore, the author believe that an artificial intuition based model can effectively contribute to the decision making and improve the overall performance of decision making process in a large network. The process bypasses some of the computational stages by simulating the process of “making assumptions”, hence it will offer considerable computational complexity savings by reducing the number of nodes to be accessed in a decision making situation.

#### *2.10. Key Requirements for the Model Development*

The studies reviewed recognize that knowledge and past experience are very important for intuition to be accurate. However, the concept of intuition, as suggested by the early studies, focuses largely on the concept itself, rather than on the representation and use of entities in the process. The earlier studies took a philosophical, psychological and cognitive approach (Kahneman, 2002; Stanovich and West, 2000; Payne, 2015; Graber et al, 2012; Hams, 2000; Kahneman, 2003; Frantz, 2003). Moreover, some of the models and frameworks identified in the studies are specifically on modelling categorical emotional states to provide direct quantitative correlation with emotional effects on decision-making (Velasquez, 1998; Sloman, 2001; Murphy et al, 2002), while others use emotion to influence cognitive processes by means of providing goal-based cues and biases. However, the effects of artificial intuition on decision-making have not been addressed. In other words, they provided some studies on artificial emotions (Salichs and Malfaz, 2012; Minsky, 2006; Rauterberg, 2010; De Melo et al, 2012; Lerner et al, 2015). To the best of our knowledge, only few studies discussed attempt to study artificial intuition from the computational point of view (Dundas and Chik, 2011; Tao and He, 2009; Srdanov et al, 2016; Diaz-Hernandez

and Gonzalez-Villela, 2017 and Liu and He, 2019). However, an important limitation of these approaches is that they did not present a detailed representation of algorithm and use of intuitive entities in the process. Moreover, Artificial Intuition still needs to be defined and modelled in order to apply it to complex decision making systems.

The review demonstrates the key capabilities of the current artificial intuition based decision making approaches. The study further highlights the merits and demerits of each of the approaches. The cross evaluation of the approaches showed there is a compelling evidence to incorporate artificial intuition models in decision making. This has motivated the conceptualization of the computational model of artificial intuition and decision making. In fact, there is a gap in this area of research. This study addresses this gap by:

- Developing a computational model of artificial intuition and decision making.
- Define and implement mathematical concepts and algorithms of artificial intuition based on network theory. More specifically, the algorithms will be driven by artificial intuition depending on the known knowledge and experience by identifying links between connected pathways that is modelled by semantic network
- Validate the algorithm
- Evaluate the model and present the results.

In this section, the authors presented some requirements of artificial intuition in decision making. These requirements are based on the literature review evidence that a computational model based on artificial intuition shall provide a better way to optimize decision making in complex decision scenarios. In fact, this study essentially identifies the presence of the following requirements.

1. Knowledge and Experience, which include prior knowledge of events, concepts, patterns as well as variables, which have been acquired over

time from experience. This knowledge of patterns has been created and is held in a mental map about the subject of interest. These patterns are relevant information pieces that help to create the intuitive decision, which is used to recognize and act rapidly. Decisions that are made in circumstances similar to previous experience, and whose outcome could be potentially harmful, or potentially advantageous, induce a somatic response used to mark future outcomes that are important to us, and to signal their danger or advantage (Damasio, 1994). In general, somatic marker is concerned about the outcome of the choices that is made about the decision options that are presented. Intuitive feeling is created based on such experiences and this aids the decision process in an automatic manner. Thus, when intuition assesses a decision context and juxtaposed a negative somatic marker to the outcome of the decision option, it sends an automatic response to drop that option.

2. Commonsense understanding and identification of subtle trends and the selective attention to certain aspects or events; the preference for viewing situations from a broader perspective.
3. The capacity for pattern recognition as well as similarity recognition, which is a comparison of similar and dissimilar characteristics; the ability to recognize subtle patterns which are relevant information pieces that enables humans to synthesize decisions. This help to create the intuitive decision
4. Have multidimensionality capacity. This means that the model will have the capability to integrate multiple threads of information simultaneously without diminishing the quality or accuracy of the response. More specifically, the system has the capacity to map or replay the patterns and adapt to uncertain and new environments.
5. Partial information about the subject of interest or context of the problem.
6. Performance in terms of the capacity to obtain answers much faster.
7. An essential aspect of the development of intuitive models is the requirement for knowledge network representation, which includes a net-



work of concepts with their corresponding properties and attributes. These concepts and their corresponding attributes are the central data structure of the model, which can be defined as semantic networks. These are a graphical knowledge representation of concepts and their mutual connections within the context of the domain that is described by the concepts. Fundamentally, the process of building a semantic network requires that documents are all restricted to a domain that can be characterized by well-defined, inter-related concepts. These concepts form the basis for the scientific terminology of the domain (Brasethvik and Gulla, 2002). Furthermore, each concept in the network is represented by a node and the hierarchical relationship between them is depicted by connecting appropriate concept nodes via is-a or instance-of links (Shastri, 1988). Nodes at the lowest level in the is-a hierarchy denote tokens, while nodes at higher levels denote classes or categories of types. Properties in the network are also represented by nodes and the fact that a property applies to a concept is represented by connecting the concept and property nodes via an appropriately labelled link.

The aim of this study is to provide a computational model of "Artificial Intuition" and decision making. This will involve a rigorous modelling of an approach that utilises semantic networks to improve a decision system. Specifically, this research hypothesizes that a computational model that correctly implement the requirements defined above can potentially obtain accurate and optimal result and improve the overall performance of human decision making systems.

### 3. Network, Graph and Semantic Network

This section introduces the relevant methods and the underpinning theory. It introduces the theory and models of networks as the suitable approaches to be used in the definition and implementation of the model in this study. It provides a discussion of semantic networks since this constitute the representation model the author have chosen to implement the model. The first section provides an overview of network and graph theory and the development of network research. The second section provides a discussion of semantic networks and some relevant examples and applications of semantic networks. The author concludes the second section by presenting a discussion of some semantic similarities and relatedness measures as well as approaches for identifying candidate terms for spreading activation in a network. This section is concluded by presenting a modeling of the evolution of semantic networks creation.

#### *3.1. Network and Graph Theory*

A network is a set of objects, called nodes or vertices that are connected together. Many of the systems in nature can be described by models of these complex networks, which are structures consisting of the connected nodes. The connections between the nodes are called edges or links. More specifically, any patterns of interactions in a given system can be represented as a network, the individual parts of the system being denoted by nodes and their interaction by edges (Newman, 2018). Numerous networks exist. For example, the Internet is a network of routers or domains. The World Wide Web (WWW) is a network of Web pages connected by hyperlinks. The brain is regarded as a network of neurons. Scientific collaboration is a network. An organization is a network of people. Human societies are networks of collections of people interacting through social relations. The global economy is a network of national economies, which are themselves networks of markets; and markets are themselves networks of interacting producers and consumers. Topics in a particular discipline can be represented as networks. From Mathematics perspective, networks are often referred to as graphs.

Network theory is the study of graphs as a representation of symmetric or asymmetric relations between discrete objects. In computer science and network science, network theory is a part of graph theory, a network can be defined as a graph in which nodes and/or edges have attributes (e.g. names). Because the task of network optimisation is *NP-hard*, the task is often broken down into subtasks by decomposing the network into relatively independent subnets (Ignatov et al, 2016).

Formally, networks consist of a collection of nodes, called the node set  $V = \{v_i\}_{i=1}^n$ , which are connected as specified by the *edge set*  $E = \{e_{w_i,j}(v_i, v_j)\}_{v_i \neq v_j \in V}$  (Albert and Barabási, 2002), excluding self-loops, that is a single edge starting and ending at the same node. We say that there is a *path*  $P(v_i, v_j)$  between the nodes  $v_i$  and  $v_j$ , if we have a sequence of edges which connect a sequence of distinct nodes, such that it starts from  $v_i$  and ends at  $v_j$ .

The topology of different networks has been extensively investigated to identify crucial information on the corresponding system, which can provide a set of predictive tools to investigate its properties (Trovati et al., 2019). In particular, stochastic topological features can successfully model systems based on unknown parameters. The networks features have been categorised into small-world, random and scale-free network (Newman, 2018). The small-world and scale-free features are common to many real-world complex networks.

### 3.1.1. Scale-free Networks

Scale-free networks can be found in a variety of contexts, in many large-scale complex networks such as the World Wide Web links, as well as biological and social networks (Humphries and Gurney, 2008; Albert and Barabási, 2002). It is important to note that as data extraction and analysis techniques continue improving, more instances of such networks will become available. The main property of scale-free networks is related to their node degrees which are governed by a power law. In other words, the fraction  $P_k$  of nodes in the network having degree  $k$  or  $k$  connections to other nodes, can be approximated, for large values of  $k$ , as:

$$pk \approx k^{-\gamma}$$

where  $\gamma$  is a parameter which has been empirically shown to be usually in the range  $2 < \gamma < 3$  (Albert and Barabási, 2002). An important feature of Scale-free network refers to the fact that new nodes are created over time, which are likely to be connected to existing nodes that are already well connected. This feature is described as the principle of “preferential attachment” (Albert and Barabási, 2002).

### 3.1.2. Random Graphs

Random graphs may be described simply by a probability distribution, or by a random process which characterises them. The main goal of the random graph theory is to determine at what connection probability  $p$  a particular property of a graph will most likely arise (Wang and Chen, 2003). In fact, such probability  $p$  specifies the existence of the edges between any two nodes. Their applicability varies across several areas in which complex networks are investigated and as a consequence, a large number of random graph models have been analysed, to address the different types of complex networks. In particular, the fraction  $pk$  of nodes with degree  $k$  follows:

$$pk \approx \frac{z^k e^z}{k!}$$

where  $z = (n - 1)p$  (Halmos, 1982), One of the crucial aspects of a random network is the fact that the set of relationships among the concepts modelled by the edges and nodes respectively, are purely random and can appear quite suddenly

### 3.1.3. Small-world Networks

The concept of the small-world network was introduced by Watts and Strogatz (1998). According to Watts and Strogatz, small-world network are a class of networks that are “highly clustered, like regular lattices, yet have small characteristic path lengths, like random graphs.”. The small world network is common, and it has received much attention due to their

applicability to a variety of real-world models. They are characterised by the property that most nodes have a relatively small degree, but they can be reached from any other node by a small number of steps. This is a popular manifestation of the “small-world effect” called the “six degrees of separation” principle (Milgram, 1967). Specifically, a small-world network is defined to be a network where the typical distance between two randomly chosen nodes grows proportionally to the logarithm. Social networks are typical examples of small-world network, in which cliques or clusters of friends being interconnected but each person is really only five or six people away from anyone else.

The discovery of small-world and scale-free properties of many natural and artificial complex networks has stimulated a great deal of interest in studying the underlying organizing principles of various complex networks and this have led to some active research in this field. Given a real network, Trovati et al. (2014) has provided an approach to determine the topological structure of the network in order to enable a full dynamical and statistical investigation of the data set(s) modelled by it. In particular, they have provided a method to assess whether the network can be approximated as either a small-world, random, or scale-free network to allow a better understanding of the associated data.

### 3.2. Graph Theory

A graph is a network of nodes and edges (connections between nodes). Each node represents a concept e.g., *Thinking* and each edge has a type e.g., *used\_for* that represents the relationship type. See Figure 3. The figure shows a network graph from ConceptNet that is formed by every activity (edges) that *Thinking* is related to. The node *idea* is created through creative “thinking”. Similarly, “thinking” is *used\_for* “solving problems”.

A directed graph is where edges link two vertices asymmetrically while undirected graphs is where edges link two vertices symmetrically.

From the mathematical perspective, graph theory is the study of mathematical properties of graphs and it provides the theoretical foundation of modern network theory (Newman, 2018).

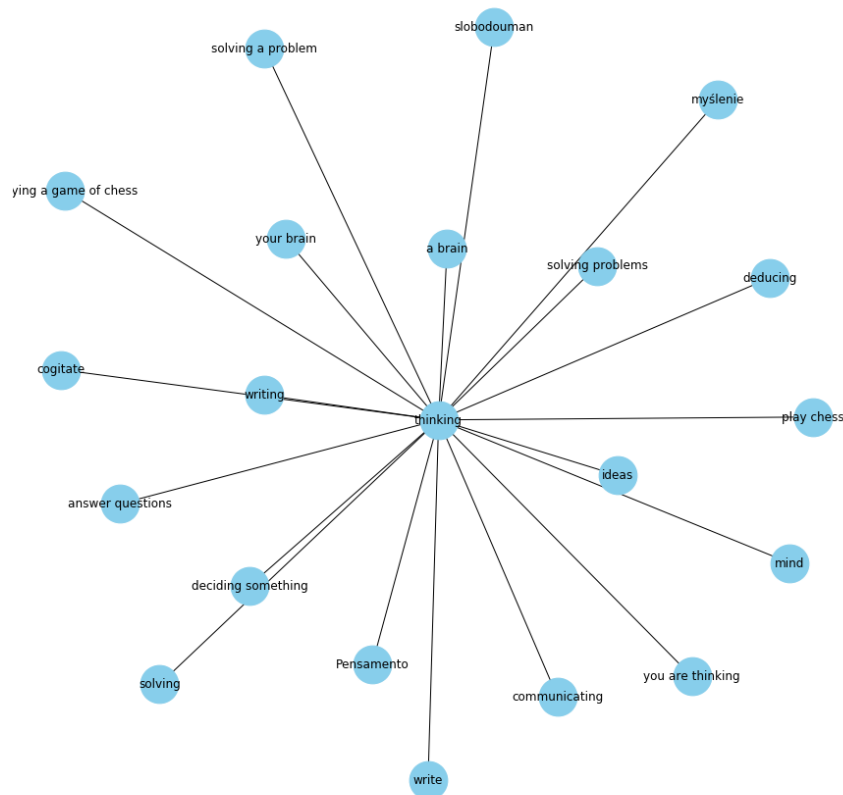


Figure 3: A network graph from ConceptNet that is formed from the concept *Thinking* and the related edges (Liu and Singh, 2004)

The great appeal of graph theory is that it is elegant, and it provides a framework to model a large set of problems in Computer science. Many Graph Problems are NP-Complete and provide a useful tool for study in Computational Complexity. Moreover, graph theory has natural connection to tools from Topology, Combinatorics and Algebra.

Networks and graph theoretical concepts are widely used to model a variety of complex, often multi-disciplinary, systems in which the relationships between their sub-parts play a significant role (Trovati et al., 2014). Network theory has increasingly attracted much interest from a variety of interdisciplinary research fields, including mathematics, computer science, biology,

and the social sciences. Their simple, yet effective formulation has allowed a successful exploitation of their applications in a wide range of real-world complex settings (Watts and Strogatz, 1998).

Moreover, computational techniques from complex network and graph theory have been used to solve classic problems about shortest path between two nodes in a graph; matching problem of pair of nodes or matching independent edge sets as well as problems of critical path analysis which is the problem of how to determine the longest path of a dependent nature in a system of interdependent activities. Moreover, quantities and measures of complex networks such as clustering coefficient and degree distribution have also been proposed and they play key roles in complex network analysis (Wang and Chen, 2003). These measurements will be explored to provide measurements of the structural properties of the network and the general network analysis in the evaluation of the model provided in this study.

### *3.3. Semantic Analysis*

Semantic analysis is a crucial part of Information Extraction (IE) in the field of Natural Language Processing (NLP) and computational linguistics. Semantic analysis deals with how the lexicons are fused together in order to make meaningful expression. Succinctly put it is the investigation of the interaction of objects associated with a lexicon and how they make meanings in the sentence.

Depending on the given scenarios and the given semantic information, a variety of NLP techniques can be used to perform semantic analysis and extract entities and relations from data sources; this however depends on the type of data and their structure. In particular, symbolic approach is one of the NLP methods for semantic analysis. It involves the investigation of linguistic phenomenon based on the explicit representation of facts about language via precise and well-understood knowledge representation. In this approach, set of rules are developed, and each rule is associated with semantic objects, which are used to validate the rule. Here concepts that are highly associated based on some rules could exhibit directly linked properties whereas weakly related concepts are linked through other semantic

objects. For example, semantic objects (concepts) associated with defined rule are more likely to exhibit properties that are directly linked. More so, rules associated with semantic objects could generate networks that describe their hierarchical structure. Symbolic methods have been exploited in a variety of research contexts such as information extraction, text categorisation, ambiguity resolution, explanation-based learning, decision trees and conceptual clustering. For example, we can apply symbolic approach during information extraction to investigate and resolve inconsistent expressions in datasets about the weather such as "Today is cloudy" and "Today is very sunny".

One of the important tasks in this study is to represent knowledge as graph. In other words, the task is to extract semantic relationships from unstructured and structured datasets. For example, Node A and Node B in figure 4 are two different entities that can be extracted from text. These nodes are connected by an edge that represents the relationship between the two nodes. This is also known as a triple. A node or an entity can also have multiple relations.

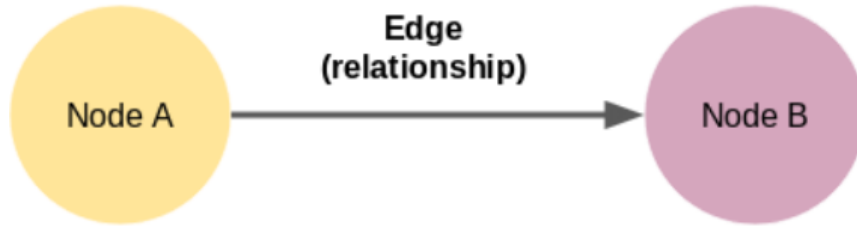


Figure 4: Graph Relationship between two Nodes extracted from texts in Wikipedia

The challenge therefore is to identify triples in the form of the type  $< \text{"Subject - verb - object"} >$  from unstructured datasets and use the triples to populate the nodes and edges of the corresponding network, by identifying any connection among the keywords defined, with the corresponding elements of the data-sets. This is by no means a trivial task. A triple represents a couple of entities and a relation between them. For example, ("a wheel is part of a car") is a triple in which *wheel* and *Car* are



the related entities, and the relation between them is "part of". Another example is "Mark Zuckerberg is the CEO of Facebook". A more complex example from spaCy (Choi et al., 2015) is:

*"Indian tennis player Sumit Nagal moved up six places from 135 to a career-best 129 in the latest men's singles ranking. The 22-year-old recently won the ATP Challenger tournament. He made his Grand Slam debut against Federer in the 2019 US Open. Nagal won the first set."*

Essentially, the task is to transform this text data into something that can be used by the machines and also can be interpreted. Some NLP approaches that can be incorporated in this are sentence segmentation, dependency parsing, parts of speech tagging, and entity recognition. This approach could involve: 1) identify the text, 2) splitting the text document into sentences and then shortlist only those sentences in which there is exactly 1 subject and 1 object, 3) Identify the entities (Entity Pairs Extraction), that is extract the subjects and the objects, 4) extract relation, which are the edges to connect the nodes (entities) to one another. These edges are the relations between a pair of nodes. 5). Finally, build the knowledge graph from the text.

By following this approach, we can extract such relations. First, for the sentence involving Mark Zuckerberg, we find the sentences with the term 'CEO' and then extract subjects and objects from those sentences. With this the entities and the relation can be identified. For this example, Mark Zuckerberg', 'Facebook' are the entities while "is" is the relation. For the second example, the sentence could be analysed, shortlisted and subsequently split as follows:

- The 22-year-old recently won ATP challenger tournament.
- Nagal won the first set.

Note that both sentences have the same relation: "won". See Figure 5 for the resulting graph of the entities and relations extracted.

Another example extracted from Wikipedia:

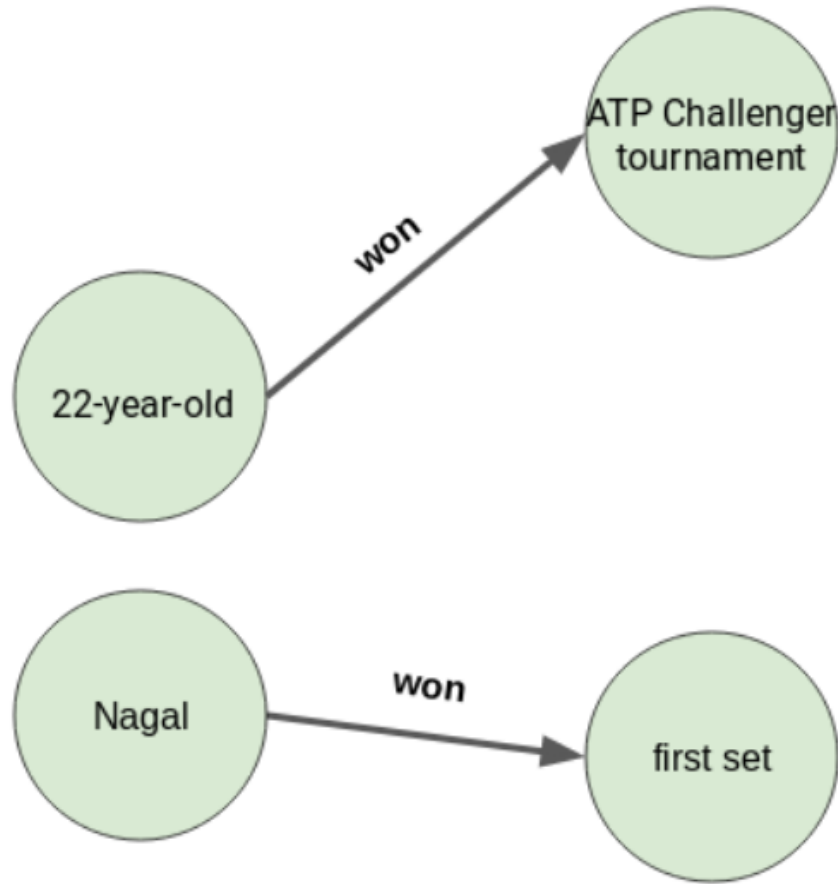


Figure 5: Example of entities and relations extracted from texts in Wikipedia

*"Food tutorial are infinitely better when directed by Wes Anderson. Bruce Lee's biopic, 'Little Dragon', is to be directed by Shekhar Kapur". "Stallone directed his first short file Vic"*

In the first sentence, there are two entities ("Food Tutorials" and "Wes Anderson"). These entities are related by the term "Directed". Hence, (Wes Anderson, directed, Food Tutorials) is a triple. Similarly, we can extract relations from the other sentences as follows:

- Wes Anderson directed Food Tutorials.
- Shekhar Kapur directed Little Dragon.

- Stallone directed Vic.

Thereafter, these extracted entities and relations are used to populate the respective nodes and edges in the network. See Figure 6 for the resulting graph of the entities and relations. Similar of such activities will be undertaken in this study to extract entities and relations from unstructured datasets.

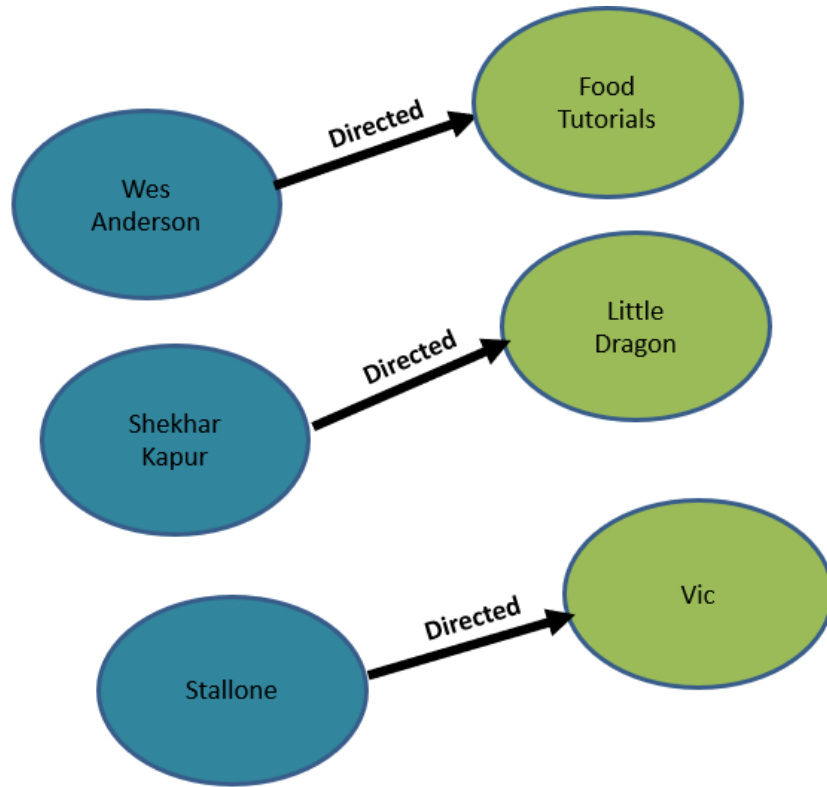


Figure 6: Example 2 of Entities and Relations Extracted

This approach can also incorporate wordnet synset. Synsets are inter-linked by means of conceptual-semantic and lexical relations. The resulting network of meaningfully related words and concepts can be navigated with the browser. Using synset will enable the identification of some structures between words that denotes concepts and some super-subordinate relation like "is-a" relation. For example, "push\_bike" and "bicycle" can also be

explained to specify a synonym, and this can be identified by using synset in wordnet.

Furthermore, this approach can also incorporate the semantic disambiguation of words with multiple senses. Semantic disambiguation allows to select the sense of ambiguous words so that they can be included in the appropriate semantic representation of the sentence. This is particularly relevant in any information retrieval and processing system based on ambiguous and partially known knowledge.

### *3.4. Semantic Network*

Since the approach presented in this study utilises semantic networks as the suitable knowledge representation, this section presents a discussion of semantic network and provide some insights and show the main properties that characterises Semantic Networks. Moreover, a discussion on the fields where semantic networks have been widely applied is presented. Finally, the author make the case for choosing semantic network as the suitable method to represent the knowledge concerning the real world in this study.

Semantic networks are specific types of graph data structure. They are graphical knowledge representation of concepts and their mutual connections within the context of the domain that is described by the concepts (Marra and Jonassen, 2002). In a semantic network, knowledge is expressed in the form of directed binary relations, represented by edges, and concepts, represented by nodes. Fundamentally, the process of building a semantic network requires that documents are all restricted to a domain that can be characterized by well-defined, inter-related concepts. These concepts then form the basis for the scientific terminology of the domain (Brasethvik and Gulla, 2002). Each concept in the network is represented by a node and the hierarchical relationship between concepts is depicted by connecting appropriate concept nodes via **is-a** or **instance-of** links (Shastri, 1988). Nodes at the lowest level in the **is-a** hierarchy denote tokens while nodes at higher levels denote classes or categories of types. Properties in the network are also represented by nodes and the fact that a property applies to a concept is represented by connecting the concept and property nodes via an

appropriately labelled link. Typically, a property is attached at the highest concept in the conceptual hierarchy to which the property applies, and if a property is attached to a node  $C$  it is assumed that it applies to all nodes that are descendants of  $C$ . In other words, all the descendants of node  $C$  inherits all the properties attached to node  $C$ . Succinctly put, subtypes inherit properties from supertypes. By suitable definition of a set of binary relations on a set of nodes, the network corresponds to a predicate logic with binary relations.

Perhaps we can present an example to show the expressive power of semantic network and how it helps to express everyday natural language. Below are some sentences that can occur in our everyday commonsense knowledge and opinion in relating to something:

1. Darren owns a Cat.
2. Cat wake\_up Henry.
3. Dami owns a Clock.
4. Clock wake\_up Temi.

Let's consider the first two sentences. We know from wordnet.synsets that the word "own" is a verb that means to "possess" something. Similarly, the word "wake\_up" is a verb that means to "awaken". Therefore, each of the above sentences are triple that represents the same type  $\langle \text{"Subject - verb - object"} \rangle$ . This is the simplest grammatical structure we can ascribe to the sentences. Moreover, the words "Darren" and "Henry" refer to names of particular persons, while the word "Cat" describes the type of mammalian. The words "owns" and "wake\_up" describes the connection between the respective persons and the animal, "Cat". With our commonsense knowledge or contextual knowledge of what the "Cat" is, we are able to understand both sentences. Note, however, that some of the knowledge can be implied. For example, the "Cat" does not directly wake up a person, rather the "Cat" makes noise and this noise in turn is capable of waking up a person. After analysing both sentences, we are able to say we can include new information to commonsense knowledge about the entire world. This is a plain example of semantics: symbols can refer to things or concepts, and

sequences of symbols express a meaning. Now, by using the meaning that we get from both of the sentences, we can answer any simple questions. For example, we can ask: "Who is the owner of this cat"? or "Who is awoken by the Cat?". Similar discussion can also be presented about sentences 3 and 4. A sequence of symbols can be used to communicate meaning, and this communication can then affect behaviour (Segaran et al., 2009).

Concepts do not exist in isolation, rather they make associations with other related concepts and this is the way by which human can potentially recall old knowledge and connect it with a new one. For example, when we read a book, we integrate the ideas expressed in the book with all that we already know, and by so doing new knowledge begins to emerge. Typical relations in a semantic network are: *Has*, *is\_a*, *is\_part\_of*, *is\_opposite\_of*, *is\_property\_of*, *is\_related\_to*, *is\_same\_as* and *is\_effect\_of*, *is\_made\_of*, *is\_type\_of*, *is\_used\_for*, *is\_caused\_by*, *is\_determined\_by*, *is\_measured\_by*. This can also be described in the form of *subject* – *>* *Predicate* – *>* *Object*. For example, *Apples* – *>* *is – a* – *>* *fruit*; *Cat* – *>* *is – a* – *>* *mammal*. In other words, relations can also be represented as a triple that has the start node, relation label, and end node: the assertion that "a dog has a tail" can be represented as (*dog*, *Has\_a*, *tail*); *IsPartOf*("a wheel is part of a car") and *Used\_for*("a car is used for driving"). From the practical point of view, triples form the fundamental building blocks of this type of knowledge representation. Each triple is composed of a subject, a predicate, and an object. For simplicity, we could think of triples just as linguistic statements from structured and unstructured datasets, where each element represent a couple of entities and a relation between them.

The types of semantic relations in a network that can be described is non-exhaustive. It depends on the scenarios and the context of use. See table 2 for a description of the semantic relations.

Many semantic networks have been developed. De Smedt (2013) presents Perception, a semantic network that stores knowledge about what things look and feel like. The database has about 9,000 manually annotated relations of the types *is-a*, *is-part-of*, *is-opposite-of*, *is-property-of*, *is-related-to*, *is-same-as* and *is-effect-of* between 4,000 mundane concepts. Relation types

Thing relation	Functionality relation	Causal relation	Event Relation
is_a	Capable_Of_Receiving_Action	Effect_Of	Subevent_Of
Part_of	capable_of	caused_by	first_subservent_of
has_a	property_of	determined_by	last_subservent_of
made_of	used_for	motivated_by	
defined_as	measured_by		
symbol_of	type_of		
opposite_of			
same_as			
related_to			
derived_from			
created_by			
similar_to			
at_location			
desires			
related_to			

Table 2: Semantic Relation Ontology

in perception are distributed across 10 contexts or categories. A portion is uncategorized. The data is visualized using nodebox. The data is available by means of an online visualizer where new relations can be added to the semantic network. The network can use a force-directed algorithm (Hellesoy and Hoover, 2006) to visualize the data.

Liu and Singh (2004) described ConceptNet, a semantic network of knowledge that consist of assertions of commonsense knowledge encompassing the spatial, physical, social, temporal, and psychological aspects of everyday life, Detail discussion of ConceptNet is presented in Section 6.1.1.

A similar Semantic Network is Wordnet (Fellbaum, 2017). WordNet is a semantic network of words. Essentially it is a database of words, primarily nouns, verbs and adjectives, organized into discrete “senses,” and linked by a small set of semantic relations such as the synonym relation and “is-a” hierarchical relations. More specifically, WordNet is based on a grouping of words into synsets or synonym rings which hold together all elements that

are considered semantically equivalent. In addition to these synset groupings. These include semantic pointers that represent relations between word meanings and lexical pointers that represent relations between word forms. WordNet is a very popular semantic network and have been applied to problems in information retrieval, data mining, and NLP. Baydin et al. (2015) have defined some correspondences between Wordnet and ConceptNet relations. For example, “canine is a hypernym of dog” in WordNet has the correspondent definition in ConceptNet as “IsA(dog, canine)”. Similarly, “automobile is a holonym of wheel” is equivalent to “PartOf(wheel, automobile)” in ConceptNet; “edible is an attribute of pear” is equivalent to “HasProperty(pear, edible)”; “to sleep is entailed by to snore” is equivalent to “Causes(sleep, snore)”. An important difference between ConceptNet and WordNet is that in ConceptNet, similarity can be done based on the concepts identified while in WordNet, similarities are done on the word matching.

BabelNet is a multilingual lexicalized semantic network and ontology (Navigli and Ponzetto, 2012). BabelNet provides concepts and named entities lexicalized in many languages and connected with large amounts of semantic relations. BabelNet was created automatically by linking Wikipedia to WordNet. The integration is done using an automatic mapping and by filling in lexical gaps in resource-poor languages by using statistical machine translation. BabelNet groups words in different languages into sets of synonyms, called Babel synsets. For each Babel synset, BabelNet provides short definitions (called glosses) in many languages harvested from both WordNet and Wikipedia. BabelNet has been applied in multilingual Natural Language Processing applications

### *3.4.1. Advantages and Disadvantages of Semantic Network*

#### **Advantages**

- It provides the expressive and representational power that enable different types of entities to be represented in Semantic Networks. for example, it enables texts from structured and unstructured data sources to be extracted and represented.



- It provides a graphical view and therefore makes it relatively easy to explore and understand the problem space.
- It provides the means to infer and derive new knowledge that are not explicitly represented in the data set. In other words, it provides the ability to work with incomplete knowledge sets representation apart from a database. This is made possible by the use of suitable and efficient inference algorithms in graph theory.
- Semantic Network can be used as a common communication and modelling tool between different fields of knowledge, e.g., between computer science and linguistic to create computational linguistics
- It provides a way to create clusters of related elements by using the appropriate semantic similarities and relatedness measures.
- Semantic Networks are characterized by a higher cognitive adequacy than logic-based formalisms.
- It presents a natural way to represent common knowledge and process information.

### **Disadvantages**

- An important issue with semantic network is that it does not make a distinction between links that constitute relation and links that are structural in nature in the network. For example, a link may have two meanings in a network, thereby causing ambiguity. See Figure 7. Brachman (1983) have identified the subtleties of the "is-a" link in a network.
- Moreover, in another case, a semantic network can have a situation where a node can point to itself, a situation commonly known as self-loop. Therefore, the meanings of links in the network is limited by the user or experts.

In these situations itemized above, the link "is\_a" behaves in a different way and have different meanings. In Figure 7, between the nodes "push\_bike" and "bicycle", it specifies an instance of "bicycle". On the other hand, the link between "push\_bike" and bicycle can also be explained to specify a synonym, this was identified by using synset in wordnet. Between the nodes "bicycle" and "transportation", the 'is-a' link can be identified to specify a category. Such knowledge can be used to organize the network in a hierarchical form. However, such problem can be resolved by making a distinction between the relational and structural links. In other words, the links between the nodes can be re-written. The link between "push\_bike" and "bicycle" can be re-written as "instance-of" and the link between "bicycle" and "transportation" can be re-written as a "subtype-of" link. See Figure 8. Similarly, the "is-a" link in Figure 9 can be explained to specify a synonym relation.



Figure 7: Example of "is-a" link that suggest ambiguity



Figure 8: Example of "is-a" link after removing the ambiguity



Figure 9: Example of "is-a" link that suggest self-loop

In both cases, additional knowledge and work are required to understand and make meaningful distinction of the links. However, we can argue that this limitation can be seen as a strength as it allows experts to have different perspectives to the same dataset that is modelled by real-world

semantic network and also the different meanings that can be derived from the resulting networks. In fact, the ability to reason over the concepts in the network provide useful modelling frameworks and an entry point into how the mind works to make intuitive decision. This is the beauty of using semantic network. The author plans to undertake similar reasonings in the Artificial Intuition modeling. In particular, cycles or self-loops, that is a single edge starting and ending at the same node will not be allowed in the implementation.

### *3.4.2. Applications of Semantic Networks*

Semantic networks have been widely applied in diverse areas of research in artificial intelligence, machine learning, cognitive sciences and reasoning (Beer, 2000; Baronchelli et al., 2013; Borge-Holthoefer and Arenas, 2010). De Deyne et al. (2016) used semantic network defined over dissimilar entities to show how association network can naturally makes correct predictions about weak similarities and the time taken to assess them. Kenett et al. (2018) used semantic network analysis to computationally examine the robustness of the memory networks of low and high creative individuals. In De Deyne et al. (2019), the authors showed an approach where they used a random walk process spreading through a semantic network to make connections between indirect concepts resulting in better prediction of human judgments of similarity than any modeling focusing alone on the two concepts at hand. Jones et al. (2015) have described semantic network as the products of experience and learning. Ferrer-i Cancho (2018) noted that no two individual semantic networks can be the same, however, there can be some consistencies due to social forces, such as communications and coordination. Wulff et al. (2018) have used semantic network to show age related difference between younger and older adults. Their study assume that individuals' semantic networks differs in both content and structure as in the case of their comparisons between younger and older adults.

The discussions showed that semantic network has a high representational and expressive power to represent knowledge as network of concepts and allow a problem space to be explored by using efficient graph algorithm.

In fact, semantic network have been key in modeling diverse phenomenon from reasoning, creativity to human cognition and decision making. In this study, the author is particularly interested in using efficient model of intuition that incorporate scenarios captured by semantic networks to improve a decision system.

### 3.5. Measurements of Semantic Similarities and Relatedness

Another important aspect to consider in our approach is the similarities and relatedness measurements of the concepts. More specifically, how do the concepts compare with respect to their semantic attributes, properties or features. Semantic similarity plays a central role in how humans process knowledge, and serves as an organization principle for classifying objects, formulating concepts, and performing generalizations and abstractions (Tversky, 1977). Therefore, it is not surprising that semantic similarity also plays a key-role in many information management tasks. Being able to effectively measure similarity is therefore a central challenge when dealing with the unstructured text.

There are fundamental differences between similarities and relatedness (Budanitsky and Hirst, 2006). Some concepts in a network can be related without being similar. For example, *Palm* and *Tree* are similar concepts and can be substituted for each other in a context. This may not necessarily be the same for semantically related concepts. For example, *ice cream* is related to spoons because we eat ice cream with spoons, but ice cream and spoons are not very similar to each other. In other words, similarities do not imply relatedness, while on the other hand, relatedness may be inferred to imply similarities. Kolb (2009) noted that dissimilar concepts can be semantically related via relations like *meronymy*(*palm-leaf*) or when concepts belong to the same semantic field (*palm-coconut*). Similarly, Turney (2008) equated relatedness with association when he noted that two words are associated when they tend to co-occur, e.g (*doctor* and *hospital*). Although this study acknowledge that some NLP applications require measures of semantic similarity, while others perform better with semantic relatedness (Sahlgren and Karlgren, 2008), however, in this study,

the author do not make such distinction between the two terms. The author only considered these distinctions out of theoretical interest and curiosity. In this study the subtle distinctions are relaxed and both terms are treated as the same. This is because measures of semantic similarity and relatedness between concepts have been widely used interchangeably in Natural Language Processing and Artificial Intelligence.

Some approaches of similarities and relatedness have been identified. Tversky (1977) showed a featural approach to similarity where concepts are compared by analysing the properties that they have in common. For example, the semantic network of *car* has a feature called *tyre* which is also a feature in the semantic network of *Bicycle*. Siblini and Kosseim (2013) use weights in weighted semantic network to measure semantic similarity. This approach is based on the idea that the types of relations that relate two concepts are a suitable indicator of the semantic relatedness between the two. Edges of the network representing the semantic relations are weighted according to the type of the semantic relation. The semantic relatedness is computed as the lowest cost path between a pair of words in the network. Thiel and Berthold (2010) showed two methods to derive two different kinds of node similarities in a network based on their neighborhood. The first similarity measure focuses on the overlap of direct and indirect neighbors while the second similarity compares nodes based on the structure of their - possibly also very distant - neighborhoods. Instead of using standard node measures, both similarities are derived from spreading activation patterns over time. In this approach, the first method of the activation patterns are directly compared, in the second method the relative change of activation over time is compared. Other approaches such as Shortest Path (Rada et al., 1989) and use of Influence levels and the change of influence level (Johnny et al., 2017) have been considered. Co-occurrence have been applied as a measure of similarity between concepts. for example, two words are associated when they tend to co-occur (*e.g, doctor* and *hospital*). In Trovati and Bessis (2016), the authors provided a method in which they used co-occurrence to analyse the topological properties of the network to determine the path-connections and the mutual influence of any two concepts corre-

sponding to specific nodes.

In this study, the choice of similarities measures is influenced by: 1) simplicity of implementation and 2) the ability to trigger pattern recognition. More specifically, if similar class of features or properties that defines two concepts can be found, it will be possible to construct an algorithm to compare them. Recognizing patterns of the features which are relevant information pieces that connect two concepts enables humans to synthesize decisions. This can potentially facilitate the process and help to create the intuitive decision making.

### 3.5.1. A Motivating Discussion

To demonstrate the similarity of concepts, let us consider an interesting discussion about the "*act of forgetting*". Some have argued that it is not only normal, but necessary, for our brain to forget material that is no longer relevant (Anderson et al., 1994). The question therefore is:

*"Is the act of forgetting a weakness or a necessary process to learn new things?."*

To answer this question, let us use the concept of a *queue* and superimpose it on the brain. The first step in solving a problem with a particular method is to determine the problem representation. The problem must be represented as a suitable form to be handled by the identified method that can potentially be used to solve the problem. More specifically, the problem of *forgetting* must be represented in a suitable form to be handled by the concept of the *queue*. That is, to simulate up to a certain point how the *brain* works during the recall of relevant information for human decision making. More specifically, we cast the problem as a *queue* problem and superimpose it on the *brain*. We can say "the brain is like the queue". The Queue resembles the *brain* in two aspects. First, knowledge is acquired by the *queue* from its environment in the form of pattern through a learning process. Second, connection strength is used to store acquired knowledge. More specifically, we define the *queue* to be analogous to the *brain*. Formally, defined as  $Brain < - > Queue$ . This is metaphorical but

life is about defining metaphors and analogies as they are used to understand things better. For more discussion about metaphors, see Lakoff and Johnson (2008).

It is known from computer science, that a *queue* is a list of items that are assessed in an ordered fashion. It is sometimes more convenient to refer to these items as knowledge. There are two components of the queue: the *least recently used* (LRU) component and the *most recently used* (MRU) component. Then, there is the process of *ageing*. The queue also has a *capacity* to hold knowledge (say 1000 items of knowledge). Facts, ideas, memories in our brain are structured as a graph, network of interconnected nodes. Some nodes are more strongly connected than others, so one can associate one fact with another faster. Now, suppose we define a semantic network for the queue:  $Part\_of(queue, LRU)$ ,  $Part\_of(queue, MRU)$ ,  $Property\_of(queue, ageing)$ ,  $Has\_a(queue, capacity)$ . Similarly, we can represent the brain using the components of the queue. In other words, we can map the domain of the queue to the brain. Therefore, their semantic networks are isomorphic as the mapping between the concepts of the two domains preserves the relationships between the concepts. That is, there is a mapping from the source domain *Queue* ( $Q$ ) to the target domain *Brain* ( $B$ ) that preserves the relationships (edges) between the concepts in  $Q$ . In the semantic networks of both  $Q$  and  $B$ , a mapping from the nodes of  $Q$  to the nodes of  $B$  preserves the edges between the concepts in  $Q$ . More specifically, this can be defined by a network  $G = G(V, E)$ , where  $V$  is the node-set containing the different concepts, and  $E$  is the edge-set so that if  $e(v_i, v_j) \in E$  then  $v_i$  and  $v_j \in V$  are assumed to be linked by an influence relation. So, if  $e(v_i, v_j) \in E$  then  $v_i$  and  $v_j \in V$  for the network  $Q$  can map to  $e(v_i, v_j) \in E$  then  $v_i$  and  $v_j \in V$  for the network  $B$ , this means that  $Q$  and  $B$  share a common relationship structure.

Similarly, to define this in terms of graph isomorphism, two graphs which contain the same number of graph vertices and connected in the same way are said to be isomorphic. Formally, two graphs  $G$  and  $H$  with graph vertices  $V_n = (1, 2, \dots, n)$  are said to be isomorphic if there is a permutation  $p$  of  $v_n$  such that  $(u, v)$  is in the set of graph edges  $E(G)$  iff  $p(u), p(v)$  is in the set

of graph edges  $E(H)$ . Although graph isomorphism is an *NP – complete* problem, meaning that an efficient structure mapping algorithm is unlikely to exist, nonetheless we can draw a similarity between the *brain* and the *queue*. Therefore, it can be argued alongside of Anderson et al. (1994) findings that the "act of forgetting is not a weakness rather it is a necessary process to learn new things".

The preservation of the relationships between concepts is an important requirement for semantic network similarity measurement. However, this preservation is relaxed in the method presented in this study for three reasons. Firstly, the author did this in order to have a more robust categories of concepts. It is sometimes more convenient to think about it in a more general term. So instead of finding a one-to-one mapping of relationships between networks, we can find more general attributes or patterns they share in common. There are various types of relations, defined by strict semantic constraints, such as temporality, type of action and direction. For example, Influence is defined by fewer constraints compared to other relations. In fact, two concepts are influenced by each other if there is a more general semantic link (Trovati et al., 2014). Loosely speaking, influence relations capture general semantic relationships suggesting a link between the corresponding concepts. In other words, the strict mathematical and semantic constraints of other types of relations, such as causal ones, are relaxed due to the vagueness inherent in influence relations. Secondly, relaxing this preservation can potentially aid an efficient computation of the similarity measurements and to find the activation value for the intuitive decision making. Thirdly, this was done in order to avoid the *NP – complete* problem. These are the factors that motivated our choice of similarity measures.

Again, going back to the example "Car reminds us of a Bicycle", "A Car is like a Bicycle". In general, the *Car* is a closely related concept to the *Bicycle* as they both have *wheels*, *tyres*, *engines* and can be defined to belong to the category of *transportation*. Similarly, the *Queue* is closely related to the *brain* as they share related properties and mode of working. So, in order to implement a similarity measure, we define a category for the concepts that have similar properties. If we can have an idea of the



attributes or features that relates concepts, we can then define and implement a function using that attribute. By using such approach, we are able to identify the class or classes related to the problem theme. As stated earlier, the human memory is a semantic network of associations where a node represents a semantic concept and these concepts are connected by varying strength of directed influence links.

The goal is to allow for association of these influence links to serve as cues to recognize patterns and act rapidly during intuitive decision-making scenarios.

### 3.6. Spreading Activation

Spreading activation (Collins and Loftus, 1975) is a method to traverse a network in order to find semantically related concepts in the network. More specifically, spreading activation is the ability to reason over a set of data based on a query and be able to find contextual neighbourhood relevant to provide the solution. Activation spreads out from the starting concept in a gradient of decreasing relatedness. The brain stores knowledge/information in the form of a graph of connected nodes/concepts and accesses the information using spreading activation and pruning. The brain builds mental links between events. An important component relevant to find the solution is the importance of the relations in the context of the domain. In the query on a semantic network, let's assume there are two important concepts that can potentially contribute to the solution domain. For example, *used\_for* and *related\_to*. if the weight of a *used\_for* relation is high (0.9) and the weight of a *related\_to* relation is low (0.1), it can be determined that the weight of the *related\_to* (0.1) reduces the strength of the connection, therefore further analysis has to be done in comparison with the other relations in the neighbourhood to find the connections whose strength is strong enough for it to be selected as part of the candidate terms or criteria to propagate or activate the solution. This processing goes on and on until the final candidate that can potentially improve the solution is selected. This is the *light bulb moment*.

Moreover, this study considers Intuition as a curious mind that searches

a path to an objective. More specifically, this study consider the curiosity as a scenario-based search space that searches the path based on the scenario. It consciously or unconsciously searches through the paths connecting the concepts that are related to each other. The more paths are searched, the more the creative idea grows and short paths to nearby concepts yield commonly available associations (Schilling, 2005). For example, “Car reminds us of a Bicycle”. Bicycle is a closely related concept to Car. In other words, both are similar and the similarity of two concepts is defined as the number of features they have in common. In fact, thinking about a *Car* or *Bicycle* reminds us of *wheel*, *tyres*, *engines*, *vehicle*, *heavy*, *travel*, *transportation*, *colour* and many more. A more relevant question to ask in this context would be, “what are the set of concepts that belong to the category of *transportation* or *travel*?”. One way to answer such question is to search for concepts that has the same semantic property with the category. Likewise, to define this relatednes, we can first hypothesize that the attributes that hold for one concept also hold for similar concepts. In fact, this generates a search process in our mind on how each of these concepts are related. However, the way each concept relates to other concepts depends on education, environment and personality (De Smedt, 2013). In fact, concepts can be related to other concepts, and this relatedness can be as deep as the representation requires within the semantic network. This process is often referred to as spreading activation (Collins and Loftus, 1975). An important aspect to consider is how do the concepts compare with respect to their semantic properties.

Spreading activation has been applied in various fields of research. Crestani (1997) applied a set of common constraints to restrict the dynamics of the process. Aswath et al. (2005) used two-level spreading activation network approach to activate strongly positive and strongly negative matches based on keyword search results. The system also used synonyms of original concepts of a query to activate, and the support vector machine (SVM) method to train and classify the positive and negative data. This is a second order term interactions derived from two-level co-occurrence data. This search was performed on a fulltext search engines. The focus of this research is on

edge activation of concepts in a network. Rocha et al. (2004) proposed a hybrid spread activation algorithm that combined spreading activation algorithm together with ontology based information retrieval. The algorithm enabled the user to represent his queries in keywords, and found concepts in the ontology having the keywords occur in their descriptions. The found concepts were counted as initial concepts and weights were assigned to links based on certain properties of the ontology, to measure the strength of the links. After that, the spreading activation algorithm was used to find related concepts in the ontology. The activated concepts did not contain any of the specified keywords. For Schumacher et al. (2008), the system found answers of given query and added into the query. After that, the system used a spreading activation algorithm to find concepts that were related to the expanded query. Berthold et al. (2009) have shown that pure (constraint-free) spreading activation with a linear activation function on a connected and not bipartite graph always converges to the principal eigenvector of the adjacency matrix of the graph. Leicht et al. (2006) presented an iterative process based on similarity measures of the nodes. Here, all the iteration results are accumulated with a decay to decrease the impact of the global node neighborhood. Thiel and Berthold (2010) argued that since the accumulated and normalized activation values are used as similarities the method may yield asymmetric similarities on directed graphs. Similarly, Pons and Latapy (2005) determined node similarities based on random walks, which are also iterative processes based. In their approach, the authors considered only the paths of a certain length when computing the similarity.

In this research approach, the author consider the semantic properties; source reliability and the probability of occurrence of the semantic properties as suitable measures of semantic activation of the edges in the semantic network which is merged from two or more sources. The author will perform the combined queries based on these conditions (semantic activation of the nodes). More importantly, the author identify the edge joining these queries in order to explore potential new solution. The author plans to achieve this by performing analysis on contextual information and use that to augment the edges from ConceptNet (the general knowledge).

To achieve the searching and comparing of relevant concepts in a dataset, it is important to consider the use of some relevant graph theoretical concepts and methods. Some relevant tools to implement spreading activation in python are *Node.flatten()* method or *cutoff* keyword argument of *nx.all\_simple\_paths* in *networkx* to generate only those paths that are shorter than a certain length. This is the termination criterion.

### 3.7. Modelling Semantic Network evolution and Decision Making

An important area in this study is how the mind conceives and evolve semantic network. How does the semantic Network evolve? The author present a brief discussion on this in the following paragraphs.

A fundamental issue that comes with the evolution of semantic network is how an algorithm can handle the requirement that the evolving semantic networks should be meaningful (Baydin et al., 2015). This is because not every node in a semantic network graph can be connected to an arbitrary other node through an arbitrary type of relation that is found in textual data. This is a relevant issue and the algorithm should be able to handle it. For examples, relations such as *is\_a(Human, Mammal)* and *caused\_by(Human, Wind)* are two relations that can be found in a text data. The difference is that the former is a meaningful relation while the latter is not. This also brings the issue of inconsistencies and noise in textual data. The issue of inconsistencies and incompleteness of data during decision making has been addressed in another section in this research.

This study addresses this issue of meaningfulness of semantic network by using commonsense knowledge and reasoning to constrain the operation of the algorithm that generates the network. Commonsense reasoning refers to the type of reasoning involved in everyday human thinking, based on commonsense knowledge that an ordinary person is expected to know, or the knowledge of how the world works (Mueller, 2014). These have been represented as semantic network. An important characteristic of a semantic network is that it is "definitional" or "assertional". In definitional networks the emphasis is on taxonomic relations, e.g. *is\_a(Bird, Animal)*; *is\_a(Human, Mammal)* and *is\_a(car, Vehicle)* describing a subsumption hierarchy that

is true by definition. In other words, we can describe this as superset and subsets; while in assertional networks, the relations describe instantiations and assertions that are contingently true e.g. *Location\_of(Car, in Garage)* and *used\_for(Alarm Clock, Wake up)* (Sowa, 1999). This study combines both approaches to expand the semantic network. The motivation for this is that it gives increased expressivity to the network. Moreover, such semantic network provides a simple way to represent the data structure. This will enable the utilization of the semantic network to evaluate our approach.

### 3.8. Algorithms For The Evolution of Semantic Networks

In the remaining part of this section, the author provides some example algorithms for the evolution of semantic networks.

#### Algorithm 1

Semantic Network of some related concepts (Encode Semantic from commonsense knowledgebase). Open Mind Commonsense have been used as a real world corpus of 400,000 facts about the everyday world.

#### Parameters:

- $size_{network}$ : Maximum size of randomly created semantic networks  $(1, \infty)$
- $score_{min}$ : Minimum quality score of commonsense relations throughout the algorithm  $(0, 1)$
- $count_{timeout}$ : Timeout value for the number of trials in commonsense retrieval operations  $(1, \infty)$

Figure 10 show a graph, which is a network of nodes and edges and it shows the process of random semantic network generation using Algorithm 1. It starts with a single random concept, *car* and adding new random concept from the commonsense database. Here, each node represents a concept (e.g., *car*) and each edge has a type of relation (e.g., *is\_a, used\_for*). Each concept activates new associative paths to follow. For example, “Car” activated “travel”. The more paths are searched, the more the idea in the

---

**Algorithm 1** Encode Semantic from Commonsense knowledgebase

---

- 1: Input: Commonsense data
  - 2: Output: Semantic Network of related concepts
  - 3: Begin
  - 4: Initialize the size of the network. ( $size_{network}, score_{min}, count_{timeout}$ )
  - 5: Identify a concept from the Commonsense data (e.g, *Car*)
  - 6: Identify list of commonsense relations associated with the identified concept (e.g. *is\_a*, *used\_for*, *part\_of*, *property\_of*, *capable\_of*)
  - 7: Identify other concepts involved in this relation.
  - 8: Append and integrate the relation and concept with the Semantic Network
  - 9: **loop**
  - 10:   Map respective influence concepts and relations to nodes
  - 11:   Repeat step 4 to 8 until the size of the network ( $size_{network}$ ) or a timeout ( $count_{timeout}$ ) is reached
  - 12: **end loop**
  - 13: Generate the network
  - 14: Validate SN
  - 15: End
- 

network grows and short paths to nearby concepts yield commonly available associations. The size of the resultant network is determined by the number of related concepts in the database or the timeout ( $count_{timeout}$ ) parameter.

**Algorithm 2**

After Algorithm 1 has been used to generate series of semantic networks, we present Algorithm 2, a mathematical algorithm to define categories and find objects with attributes in those categories in a semantic network.

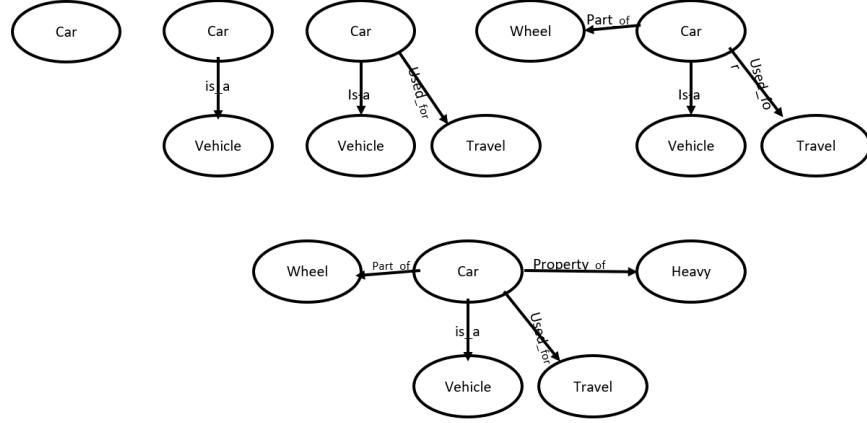


Figure 10: Evolution of Semantic Network using algorithm 1. This is an example of a small-world network

---

**Algorithm 2** Mathematical Algorithm to define categories and find objects

---

- 1: Begin
  - 2: Get search keyword
  - 3: Get category (*keyword*)
  - 4: Define category ( $C_1$ ) :  $X_{ci} = \{x_i : c_i \neq \phi\}$
  - 5: Define Objects= $x_i \in V : X_i = \{X_1, X_2, X_3, X_4 \dots X_n\}$
  - 6: Expand concepts(keyword)
  - 7: Find objects with attributes in those categories
  - 8: **for** m category associated with keyword **do**
  - 9:   find  $\prod_{j=1}^m X_{ci} := Results$
  - 10: **end for**
  - 11: End
- 

**Algorithm 3: Ranking of Concepts in Result and finding the activation value.**

After the concepts relatedness in their respective categories have been identified in the semantic network, the next important task is to rank the elements and find the activation value with respect to the main solution class or classes to activate a possible solution. The activation is influenced or determined by prior knowledge and experience. The relevant question to ask here is how each of the concepts influences the main solution space. To answer this question, a solution could be to consider the following possible

methods of activation:

- Using Influence levels
- Using the lowest cost path between classes in the network.

The mathematical algorithm presented here ranks the concepts in the categories and find the activation value with respect to main solution class or classes to activate a possible solution. This algorithm proceeds from Algorithm 1 after the influence levels or scores of the respective concepts have been calculated and assigned.

---

**Algorithm 3** Ranking concepts and finding activation value

---

```

1: Begin
2: Get search keyword
3: Get category (keyword)
4: Define category ( $C_1$ ) :  $X_{ci} = \{x_i : c_i \neq \phi\}$ 
5: Define Objects= $x_i \in V : X_i = \{X_1, X_2, X_3, X_4 \dots X_n\}$ 
6: Expand concepts (keyword)
7: Find objects with attributes in those categories
8: for  $m$  category associated with keyword, find  $\prod_{j=1}^m X_{c_i} := Results$  do
9:   for each  $r_i$  in results do
10:     Obtain the influence level ( $\tilde{I}(A, B)$ )
11:     if Score, ( $\tilde{I}(A, B)$ )  $\leq Score_{min}$ , (that means score is close to 0)
12:       then
13:         add  $r$  to solution set
14:       end if
15:   end for
16: end for
17: End

```

---

### 3.9. Summary

In the first part of this chapter, the author have presented a discussion of data and knowledge inconsistencies and incompleteness and it implication for decision making. It discussed how the issues of inconsistencies and incompleteness of data can be addressed during the extraction of networks



from datasets. The second part introduced and defined the theory and models of network, graphs and semantic networks as the suitable model to be used in this study. Moreover, a discussion of semantic similarities and relatedness measures as well as approaches for identifying candidate terms for spreading activation in a network was presented. This chapter was concluded by presenting algorithms for a modeling of the evolution of semantic networks creation. In the next chapter, a model of gut-feeling that utilises semantic network is introduced and discussed.

## 4. Gut-feeling in Decision Making: An initial Approach for Influence Relation Assessment

In Section 2.6, a discussion of gut-feeling in the decision making process was presented. This chapter provides a description of a novel approach for assessing influence relationships between pairs of concepts extracted from textual sources. The approach is based on an interpretation of the “gut-feeling”, which provides an accurate and computationally efficient approach to the decision-making process. The approach described is in line with the findings that gut-feeling influences human decision making process in complex scenarios. This approach is provided as part of a wider line of inquiry into artificial intuition decision processes.

The approach described is in line with the findings that gut-feeling influences human decision making process in complex scenarios. Although, this study is not about modelling categorical emotion and decision making, the model assume that human processes are susceptible to the effects of emotional state. The justification for this modeling assumption is that there is a resonance between the emotional content of memory and the overall emotional state and this act as a cue to gut-feeling. This is part of the associative-semantic network. More specifically, the emotional state of a person is one of the components that feeds into the gut-feeling decision making process.

In modelling a complex scenario, one of the first tasks is to assess the existence of relationships linking its different components. There are various types of relations, defined by strict semantic constraints, such as temporality, type of action and direction. Influence is defined by fewer constraints compared to other relations. In fact, two concepts are influenced by each other if there is a more general semantic link (Trovati et al., 2015). Loosely speaking, influence relations capture general semantic relationships suggesting a link between the corresponding concepts. In other words, the strict mathematical and semantic constraints of other types of relations, such as causal ones, are relaxed due to the vagueness inherent in influence relations.

The aim of this section is to provide an implementation, which models

a “gut-feeling” approach in the assessment of an influence relation between pairs of concepts extracted from textual sources. The proposed model comprises of the following essential components (assumptions):

- Some initial knowledge regarding the influence relationships between the corresponding concepts is present
- Additional information is then obtained iteratively via appropriate text analysis of small sets of texts.
- If the overall knowledge is consistent with the initial assessment, then it is assumed it is accurate and no further analysis is suggested. This suggests that the gut-feeling associated with the original “shallow knowledge” (as opposed to a deep analysis of the necessary data) is an appropriate modelling representation of the corresponding scenario.

There are a variety of approaches to extract and evaluate relevant information to assess influence relations between concepts (Shachter and Bhattacharjya, 2010). All these methods are based on rational thinking, which virtually assumes unlimited knowledge, time, and information-processing power. In other words, all possible scenarios are considered, and their outcomes are assessed via a logical and systematic manner to identify the best possible choice. The ability to perform such evaluation based on large quantities of parameters, which define a specific scenario is at the core of deep learning. However, there is compelling evidence from neuroscience research that emotions play an important role in the rational process of intelligence and the decision-making process (Loewenstein and Lerner, 2003).

Emotions are changes in both body and brain states in response to different stimuli (Damasio, 1994). When physiological changes occur in the body, they are relayed to the brain where they are transformed into an emotion. In fact, over time, these emotions and their corresponding states of the body become associated with particular situations and their past outcomes. Consequently, when making decision, these physiological changes and their evoked emotion are consciously or unconsciously associated with their past outcomes, which therefore influence the overall decision-making

process. This reasoning and decision-making is said to be associative because it compares similar situations that have been encountered in the past with current situations and then make decision accordingly. On the other hand, some have argued that rational thinking and decision-making does not leave much room for emotions (Livet, 2010). In Zeelenberg et al. (2008), the authors cautioned that we are only rational within the limits of our cognitive capacities and that decision making itself is often an emotional process, and without emotional involvement, decision making might not even be possible or might be far from optimal (Damasio, 1994). In fact, the decision making process depends on emotional processing and the resulting feelings, which involve images that relate to the state of the body. This study agrees with the views of Damasio (1994) and Zeelenberg et al. (2008), captured by the following statement: “reason without emotion is inadequate for making the decisions that guide our lives, and in fact make up our lives” (Binali et al., 2010). This provides motivation and a new perspective on incorporating gut-feeling in decision-making systems. Furthermore, an automated method based on a “gut-feeling” assessment would enable a more computationally efficient approach in this context.

#### 4.1. Description of the Approach

As discussed above, the main intuition behind the method proposed in this section is that when we “feel” that a specific decision or information assessment is to be preferred over other options, this is based on some background knowledge. Usually, this type of knowledge has a variety of sources, from personal experience, to common knowledge and beliefs. An exhaustive modelling of all the possible types of knowledge, which play a central role in this aspect is beyond the scope of this study as it would require a full investigation of several psychological and sociological issues. As a consequence, this study assume the existence of some *a priori knowledge* in the form of structured data where concepts are linked by mutual influence relations, which vary in strength. In particular, this can be defined by a network  $G = G(V, E)$ , where  $V$  is the node-set containing the different concepts, and  $E$  is the edge-set so that if  $e(v_i, v_j) \in E$  then  $v_i$  and  $v_j \in V$  are assumed to

be linked by an influence relation (Trovati and Bessis, 2016). Note that influence can be viewed as a many-to-one map. Figure 11 depict the influence relation ( $e_{v_i, v_j} \in E$  and  $v_i, v_j \in V$ )

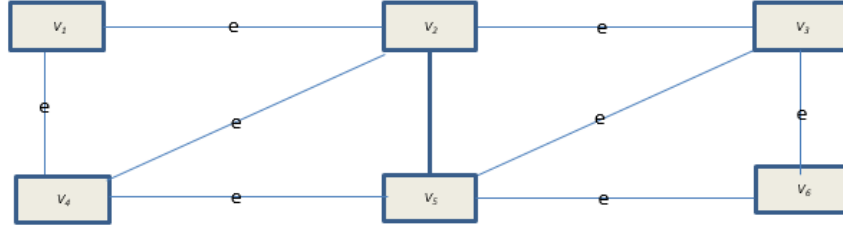


Figure 11: Influence Relation/Link in a Network

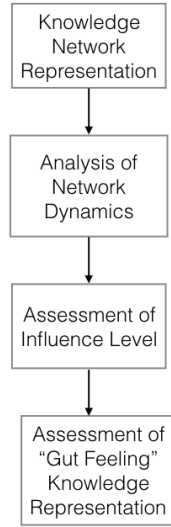


Figure 12: The Components of Gut-feeling Method

As depicted in Figure 12, the dynamical properties captured by the network are investigated to provide an assessment of the reliability of the initial knowledge on a specific scenario. In particular, we defined a knowledge set as stable if its properties remain (relatively) constant. This allows us to

measure the level of reliability of some *existing knowledge* when it used to infer an outcome. In other words, let  $G_t = G_t(V_t, E_t)$  be the knowledge at a time  $t$ . If the dynamics of  $G_t$  is stable over a time iteration, then it is assumed it can be reliably used. Consider the following example which suggest a *prior knowledge* of the dynamics of the rain.

- If it is windy and cloudy in the morning, it rains
- If it is windy and cloudy in the afternoon it rains
- If it is windy and cloudy in the evening it rains.

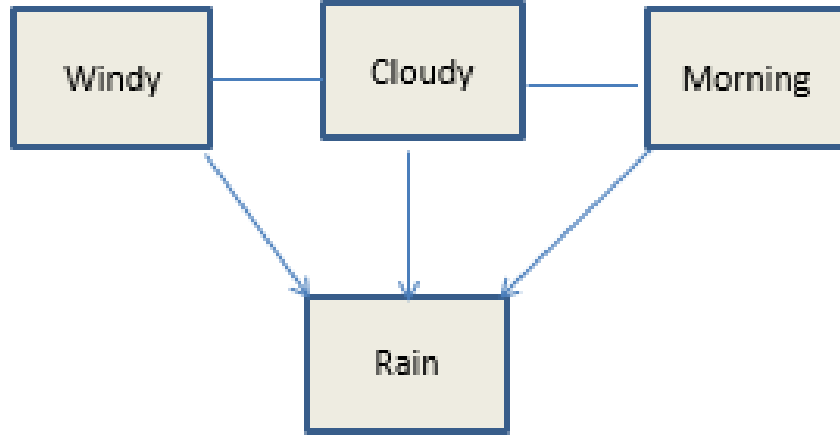


Figure 13: Network of Influence link between some *prior knowledge* of the dynamics of the rain

Figure 14 captures the knowledge of the rain at a particular time in the past (say *2pm* yesterday).

For  $e_{v_i, v_j} \in E$ , we define the influence level as  $\tilde{I}(e_{v_i, v_j}) \in (0, 1]$ , so that if  $\tilde{I}(e_{v_i, v_j}) = 1$ , then  $v_i$  and  $v_j$  are strongly influenced by one another. Note that influence relations might not have a clear direction, as they are assumed to be much more general than directed relationships such as causality (Trovati, 2015). Loosely speaking, if two concepts influence each other, then we know they are linked without necessarily knowing which of them is directly affected by the other. In order to assess the dynamics of  $G_t$  for

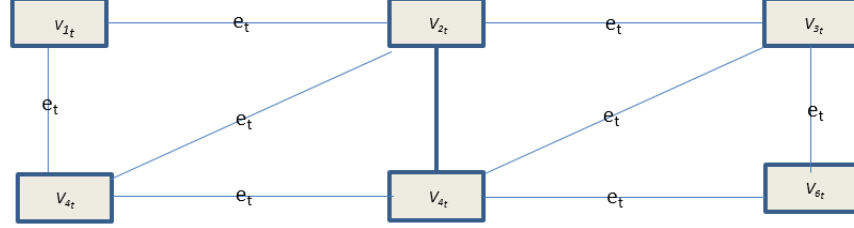


Figure 14: Knowledge of the rain at a particular time in the past. It captures the reliability of the influence link over a period of time. This measures the entropy and the influence link in making a decision in the network

$t \geq 1$ , we need to evaluate the differences between  $G_t$  and  $G_T$  for  $T > t$ . We assume that no nodes can be removed, or in other words,  $V_t \subset V_T$ . However, new nodes can be added.

After obtaining the knowledge, the next step is to test for reliability of the properties of the initial (existing) knowledge. How reliable is this initial knowledge over a certain time iteration?. *Is the knowledge stable and consistent? can the knowledge be trusted in a specific scenario or across all scenarios over time?.* if the properties are reliable over time then the initial (existing) knowledge can be used. An approach to test for reliability over a period of time is to apply entropy, Entropy is a measure of knowledge disorder.

In assessing the reliability of our current knowledge regarding the influence between two concepts  $A$  and  $B$ , the following will be considered.

- The shortest path connecting  $A$  and  $B$  with the biggest average influence value over its edges. If we have more than one, any of them can be chosen
- The average change of the influence between concepts in the path between  $A$  and  $B$ , and.
- How widely they change over a certain amount of time iterations.

The former can be modelled via the following equation

$$\frac{1}{m} \left| \sum_{k=2}^m \left( \sum_{e_{v_i, v_j} \in p(A, B)} \frac{1}{\#E_{p(A, B)}} |i_1(e_{v_i, v_j}) - i_k(e_{v_i, v_j})| \right) \right|, \quad (1)$$

where  $p(A, B)$  is the path connecting  $A$  and  $B$ ,  $\#E_{p(A, B)}$  is the number of edges in the path and  $i_l(e_{v_i, v_j})$  is the influence level at time  $l$ .

The latter is modelled via the following equation

$$\log \left( \frac{n}{W} \right), \quad (2)$$

where  $W = \#\{w : |i_1(e_{v_i, v_j}) - i_k(e_{v_i, v_j})| > th, k = 2, \dots, m\}$  and  $th$  is the outlier threshold of the influence, which can be either set manually, or inferred from a training dataset. In other words,  $W$  measures the number of influence values far from the initial value at  $t = 1$ , relatively to the threshold  $th$ .  $n$  is the number of edges.

Therefore, we define the *change of influence level*  $\tilde{i}(A, B) \in [0, 1]$  by combining Equations 1 and 2

$$\begin{aligned} \tilde{i}(A, B) = & \\ \frac{1}{2} \left( \frac{1}{m} \left| \sum_{k=2}^m \left( \sum_{e_{v_i, v_j} \in p(A, B)} \frac{1}{\#E_{p(A, B)}} |i_1(e_{v_i, v_j}) - i_k(e_{v_i, v_j})| \right) \right| \right. & (3) \\ & \left. + \log \left( \frac{n}{W} \right) \right) \end{aligned}$$

Note that if  $\tilde{i}(A, B)$  is close to 0, then the initial evaluation of the influence relation level is reliable. On the other hand, if it is close to 1, then it should be discarded suggesting that a deeper analysis should be carried out.

A preliminary evaluation of the method described above was done based an annotated textual dataset extracted from PubMed (Canese and Weis, 2013) which includes over 24 million citations from the biomedical research field. Three keywords pairs, which correspond to well known and widely documented concepts that are influenced by each other were selected. The



validation was done by considering the influence relationships between the concepts based on an interpretation of their threshold and Influence levels. The result suggest a promise for a computationally efficient approach to the decision making process.

#### *4.2. Summary*

In this chapter, the author have provided a preliminary approach to assess influence relationships between concepts and a scenario for its use. This is part of a wider line of inquiry into artificial intuition and decision making. The aims are to define complex decisional networks by bypassing complex calculations by providing an innovative approach to their analysis. In the next chapters, this study will be providing a more rigorous modeling of artificial intuition in specific decision making scenarios as well as a better predictive assessment of the artificial intuition levels to be embedded in the creation of suitable semantic networks.

## 5. The Proposed Solution

In this chapter, a rigorous approach to model artificial intuition is proposed. The aim is to facilitate a comprehensive theory and subsequent implementation of Artificial Intuition to provide a better decision system, which mimics the agile and efficient intuitive processes extensively used by human agents.

### 5.1. *Defining Essential Components of the Model*

In Johnny et al. (2019), the authors presented an initial discussion on the main architecture and some requirements of artificial intuition in decision making. See Figures 15 and 16. As depicted in the figures, the overall approach is the decomposition and analysis of the objectives within a specific scenario. In particular,

1. An objective is defined as a collection of semantic properties and specific features.
2. Curiosity is the process to discover new knowledge.
3. “random curiosity” is searching or observing without clear objectives.
4. “targeted curiosity” (which is what we consider here, and we will refer to it as simply “curiosity”) is based on specific objectives and we are looking for any potential new way to do things.
5. In any such process, there is a waste of energy or cost involved.
6. Therefore, there will be a trade-off stage, which will need to be addressed in future work.

In this section, a method which models an “artificial intuition” approach that utilises semantic networks to improve a decision system is introduced.

The author define artificial intuition as the ability of a system to assess a problem context and use pattern recognition or properties from a dataset to choose a course of action or aid the decision process in an automatic manner. Essentially, modeling of artificial intuition is accomplished through recognition of significant patterns and properties that are made available by prior knowledge and experience, given the context of the problem. More

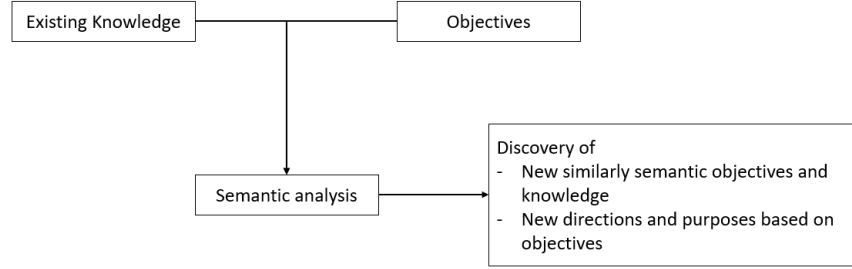


Figure 15: Modelling Curiosity

specifically, by using intuitive models, a system is able to take subsets from networks and pass them through a process to determine relationship that can be used to predict future decision without a deep understanding of a scenario and its corresponding parameters. The proposed model is based on the following essential components.

1. Semantic network representation of existing and commonsense knowledge: Some initial and existing knowledge regarding the relationships between the corresponding concepts within a specific setting is present and captured in a semantic network. These relationships are the significant patterns and properties as well as variables, which have been acquired over time from experience. This is the central data structure that have been created and held in the mental model about the subject of interest. In other words, these are associative-semantic network that forms its long-term memory (Bower and Cohen, 2014) featuring the conceptual nodes and the associated links. Bower and Cohen (2014) represents the human memory as a semantic network of associations where a node represents a semantic concept and these concepts are connected by directed links. The strength of these influence links varies, and these links allow for association and serves as cues to recognize and act rapidly during intuitive decision making scenarios.
2. Semantic network associated with the contextualised knowledge.
3. Analysis of the network dynamics: Here the dynamical properties captured by the networks are investigated to provide an assessment of the

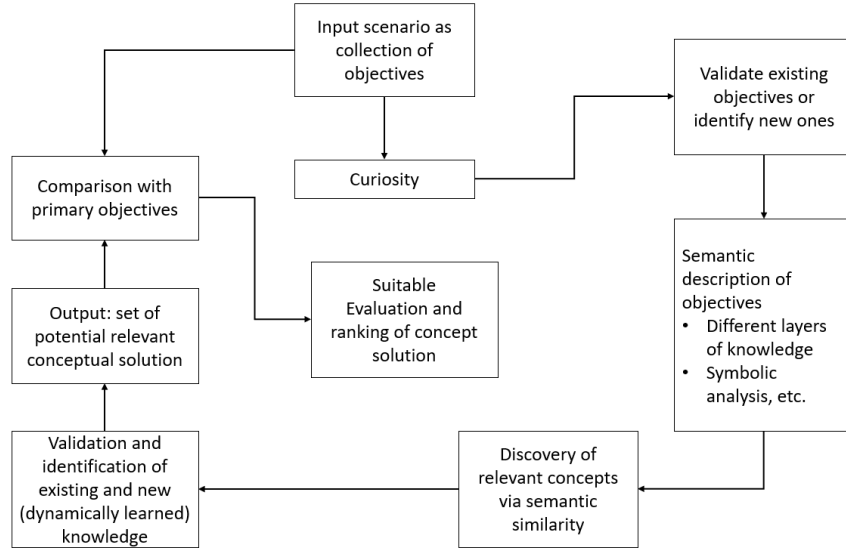


Figure 16: The Main Components of the Method

reliability of the initial knowledge on a specific scenario. Additional information is then obtained iteratively via appropriate network analysis of subset of the networks.

4. Assessment of the concepts' relatedness in their respective categories and similarity measurements.
5. Assessment of the intuitive decision making process: If the overall knowledge is consistent with the initial assessment, then we assume it is accurate and no further analysis is suggested.

The author shall now move to the next section to provide the description of the techniques that model the approach.

### 5.2. General Definitions and Background of Main model

The main motivation behind the design of a possible algorithm stems from the simple observation that intuition identifies new pathways between a starting set of concepts, to a specific target. More specifically, intuition can be used to discover new connections and pathways to identify potential

solutions. In this context, this process will be defined as a *query* as described in Definition 1.

**Definition 1.** *A query is a collection of semantically-linked concepts, which is the main scenario or objective to be modelled. The solution of a query is a set of paths linking the output with suitably identified concepts. The concepts at the end of a path related to a query are called leaf nodes.*

To introduce the main idea, consider the following simple example, as depicted in Figure 17. Suppose the following query “I need to wake up at 7am tomorrow” needs to be addressed. The obvious step is to set up my alarm clock. Assume its batteries are almost flat. However, I have just realised that the milkman delivers my allocated milk at exactly 7am. Moreover, my cat, who is very susceptible to noise, will react to hearing the milkman’s arriving in his van by jumping on my bed. Therefore, I could use this chain of events namely: “*milkman – van – engine – noise – cat*” to identify the solution of this query. Essentially the solution is activated by using the integration of the semantic properties of the “*Cat*” and “*Milkman*” to find the potential solution to “wake up”. The semantic properties of the Cat (Breaths, makes noise, hear, purr), The Milkman (delivers milk, drives van), the van (engine, wheel), engine (makes noise), Cat (react to noise, make noise). Essentially, we have integrated the concepts using their semantic properties in order to arrive at the potential solution.

In other words, the above process can also be viewed as depicted in Figures 18(a) and 18(b), where the former represents a conventional discovery approach, whereas the latter represent how intuition can be used to make the overall discovery process more efficient.

The model proposed in this study will be based on three different types of knowledge as introduced in Definition 2.

**Definition 2.** *We define*

- Existing knowledge as information associated with specific and well-known knowledge

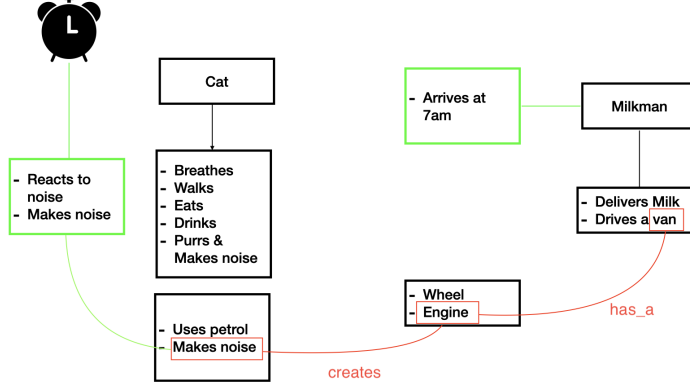


Figure 17: The example discussed in Section 5.2.

- Intuitive knowledge *as information associated with more general knowledge, which might complement the above*
- Contextualised knowledge *as information associated with individual experience and knowledge, if applicable.*

As a consequence, it is not surprising that network and graph theory is likely to be one of the most suitable approaches in this endeavour. A rigorous description of graphs has been presented and discussed in chapter 3. In this section, the main mathematical concepts and algorithms, which will be evaluated and assessed in Chapter 6 shall be described.

In this work, the author defined an undirected network  $G = G(V, E)$ , where  $V = \{v_i\}_{i=1}^n$  is the *node set* and  $E = \{e_{w_{i,j}}(v_i, v_j)\}_{v_i \neq v_j \in V}$  is the *edge set*. Note that each edge  $e_{w_{i,j}}(v_i, v_j)$  is weighted by the parameter  $w_{i,j} \in (0, 1]$ , which is related to the type of relationship linking the two nodes  $v_i$  and  $v_j$ . We say that two nodes are *adjacent* if they are connected by an edge, and two edges are *incident* if they have a node in common. We define a *path*  $P(v_a, v_b)$  between two nodes  $v_a$  and  $v_b$  a sequences of incident edges

$$e(v_a, v_{k_1}), e(v_{k_1}, v_{k_2}), \dots, e(v_{k_{n-1}}, v_{k_n}), e(v_{k_n}, v_b)$$

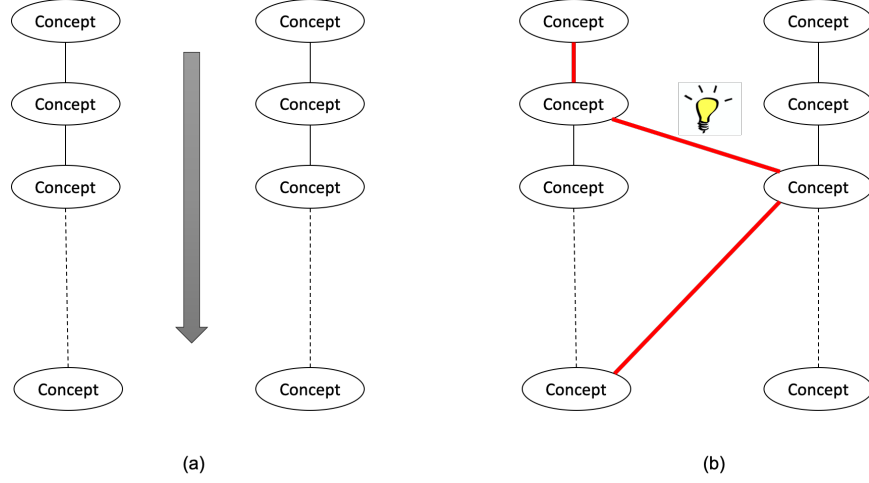


Figure 18: A general depiction of artificial intuition process as the ‘lightbulb’ moment.

joining the two nodes. Note that if a network is not acyclic, then more than a path might exist between any two nodes.

In this study, as discussed in Definition 2, we shall consider the network generated by the union of three (usually overlapping) following networks

$$G = G_k \cup G_i \cup G_c \quad (4)$$

where

- $G_k$  is the (semantic) network associated with the existing knowledge within a specific setting,
- $G_i$  is the (semantic) network associated with the intuitive knowledge and
- $G_c$  is the (semantic) network associated with the contextualised knowledge.

Loosely speaking,  $G_k$  is associated with specific, well-known knowledge;  $G_i$  refers to a more general knowledge, which might complement  $G_k$ ; finally,  $G_c$  contains knowledge related to individual experience and knowledge.

As discussed above, each node is associated with a specific concept and the overall topology of the network  $G$  governs the way information is propagated across the network. However, another crucial aspect is the way information is propagated across the network. More specifically, the overall information captured by the system comprises

- The (perceived) probability of occurrence of each concept,
- The types of relationships associated with each edge, and
- The *influence weight* of each of them.

Loosely speaking, the influence weight refers to the “strength” of the corresponding relation, which will also characterise how information is propagated along each associated edge. The above points will be discussed in the rest of the section.

### 5.3. Description of the Main Artificial Intuition Model

In this section, a model for artificial intuition is proposed and discussed. The majority of the properties of the model are inherited from network theory, since the mutual interactions within knowledge systems can be efficiently described as networks, where concepts correspond to nodes, linked by suitably defined edges, containing all the relevant information on the relationships joining any two nodes. More formally, each edge joining two nodes  $x, y \in V$  has an associated *activation value*  $\alpha(x, y) \in (0, 1]$  which is associated with the influence that  $x$  exerts on  $y$ .

Edges link nodes based on their semantic properties. Recall that paths consist of a sequence of incident edges (and so adjacent nodes) joining a start node with an end node. The general architecture of the approach introduced in this study is depicted in Figure 19. More specifically, the concepts and their mutual relations within a query are identified either manually or via automated methods. In this study, only the former will be considered. This will create a semantic network which will be embedded into a suitably defined network generated by the corresponding existing, intuitive and contextualised knowledge. The topology of such network will be used to identify novel and potentially innovative solutions to a given query.



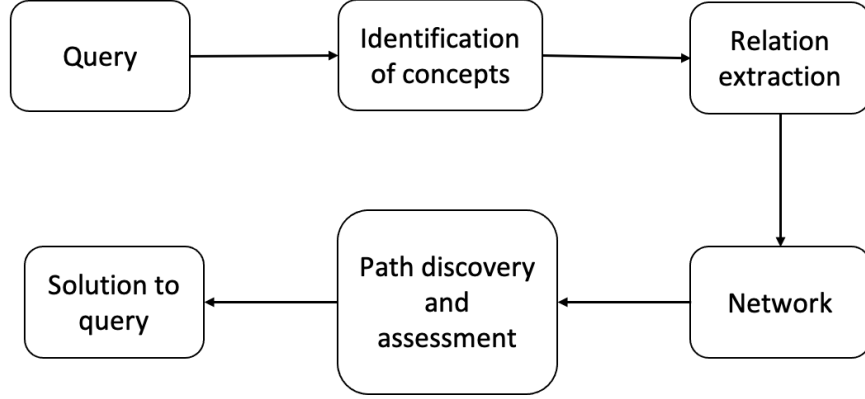


Figure 19: Overview of general components of the model

**Definition 3.** *The information propagation  $I(x, y)$  between two nodes  $x$  and  $y$  is defined as a map*

$$I : V \times V \rightarrow [0, 1]. \quad (5)$$

$x, y \in V$ ,  $e(x, y) \in E$ ,  $p(x)$  is the probability of the node  $x$ , and  $i(e(x, y)) \in (0, 1]$  is the *influence weight* of the edge  $e(x, y)$ . An important observation is that the information propagation from  $x$  to  $y$  might not necessarily coincide with the observed probability of  $y$ . Therefore, we define the *post node probability* as

$$\tilde{p}(v) = \min \{I(v), p(v)\}, \quad (6)$$

where  $p(v)$  is the observed probability of  $v$ .

**Proposition 1.** *Let  $P(x_a, x_b)$  be a path between  $x_a$  and  $x_b$ . Recall that the set of neighbours of a node  $z$  is denoted as  $n(z)$ . Let*

$$\tilde{V} = \{z : z \in P(x_a, x_b) \cup n(z)\}.$$

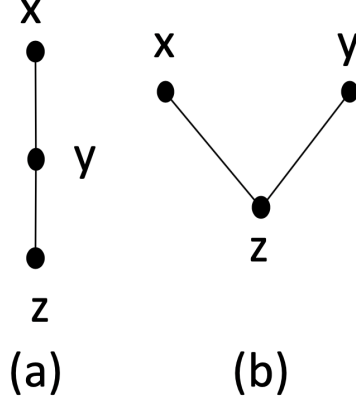


Figure 20: Two simple networks as discussed in Section 5.3.1.

We then have that

$$I(x_b) \leq \prod_{v_i \in \tilde{V}} p(x_i) \prod_{x_i \neq x_j \in \tilde{V}_l} (1 - \exp(w(e(x, y)))) \quad (7)$$

$$I(x_b) \geq \prod_{v_i \in \tilde{V}} p(x_i) \prod_{x_i \neq x_j \in \tilde{V}} (1 - \exp(w(e(x, y)))) \quad (8)$$

where  $\tilde{V}_l \subset \tilde{V}$  is the set of leaf-nodes.

*Proof.* It follows from the above definitions. □

### 5.3.1. Combination of Edge Attributes

The information propagated along the edges feeds into the nodes in the corresponding paths. Consider, for example Figure 20 which depicts two very simple networks consisting of 3 nodes, namely  $x, y$  and  $z$ . These simple configurations might be associated with the following possibilities:

- Both  $x$  and  $y$  directly influence  $z$ , as depicted by Figure 20(a). In other words, they need to co-exist in order to have  $z$ . Consider for example Figure 21, where both *wings* and *propulsion* are necessary to

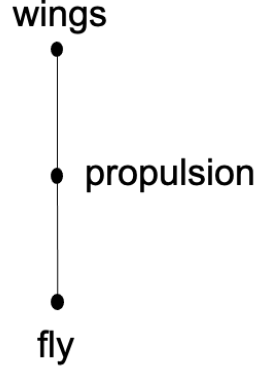


Figure 21: The “wings–propulsion–fly” network.

achieve the state *fly*. If either of them is not present (or too “weak”) then flying might not be achievable.

- Figure 20(b) depicts a cumulatively influence of  $x$  and  $y$  on  $z$ . The network depicted in Figure 22 describes a simple scenario where both the engine and the wheels (driving on the road) of a car produce noise. However, the overall noise consists of the cumulative combination of them. In other words, even if any of the two is not present, information is still propagated.
- Figure 23 has the identical topology as Figure 22. However, this scenario is completely different as it only captures semantic relationships. In fact, despite both *dog* and *fish* being connected to *animal* (as they both are), the state of one of them does not influence the other one. In such case, they would be considered independent.

The above cases are formally written as

$$x \oplus y \rightarrow z \quad (9)$$

$$x \odot y \rightarrow z \quad (10)$$

$$x \not\leftrightarrow y \rightarrow z, \quad (11)$$

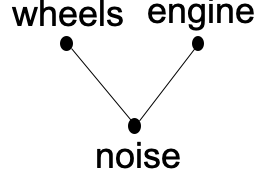


Figure 22: The “wheels–engine–noise” network.

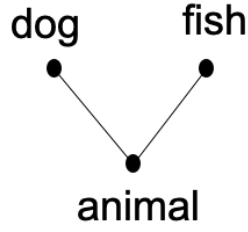


Figure 23: The “dog–fish–animal” network.

where ‘ $\oplus$ ’, ‘ $\odot$ ’ and ‘ $\not\rightarrow$ ’ refer to the disjoint, joint as independence relationships, as per Figures 21, 22 and 23. Note that the above expressions also include ‘ $\rightarrow z$ ’. This refers (with a slight abuse of notation) to the fact that we are considering the influence of the nodes  $x$  and  $y$  have on the node  $z$ . This notation will be dropped when such influence does not need to be emphasised. Consider the simple example depicted in Figure 24. whose relationships can be described as

$$(x_1 \odot x_4 \odot x_5) \oplus (x_1 \odot x_2 \odot x_3) \rightarrow x_6.$$

**Lemma 1.** *Let  $x$  and  $y$  be two nodes. Then we have that*

$$x \oplus y \equiv x \odot y \Leftrightarrow x \not\rightarrow y.$$

*Proof.* This is can be easily observed from the above. In fact, the disjoint and joint relationships yields the the same influence on, say  $z$ , then clearly

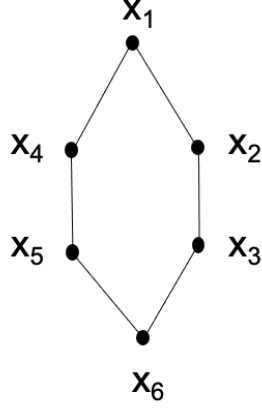


Figure 24: A simple example.

they are independent. □

### 5.3.2. Explicit Calculation of the Node Combinations

Since the activation value associated with each edge joining two nodes  $x$  and  $y$  governs the information propagation, it will be an important parameter in the explicit formulation of Equations 9 and 10.

More specifically, for three nodes  $x, y, z$  we define  $x \odot y(\rightarrow z)$  as

$$x \odot y = p(x)p(y)W_{x,y}W_{y,z}, \quad (12)$$

where

$$W_{x,y} = \tanh\left(\frac{k \alpha(x,y)}{2}\right), \quad (13)$$

where the choice of  $k$  depends on how steep  $W_{x,y}$  needs to be. Equation 13 is motivated by the sigmoid activation function widely used in Artificial Neural Networks Wanto et al. (2017).

Similarly, the disjoint relationship operation  $x \oplus y \rightarrow z$  as

$$x \oplus y = \begin{cases} p(x)W_{x,z} + p(y)W_{y,z}, & \text{if } p(x)W_{x,z} + p(y)W_{y,z} \leq 1 \\ 1, & \text{otherwise.} \end{cases} \quad (14)$$

**Lemma 2.** *Let  $x, y$  and  $z$  be nodes of a network as defined above. Therefore we have that the disjoint and joint relationships are*

- *Commutative, that is  $x \odot y = y \odot x$  and  $x \oplus y = y \oplus x$ ,*
- *Distributive with respect to  $\oplus$ , that is  $(x \oplus y) \odot z = (x \odot z) \oplus (y \odot z)$*
- *Non-distributive with respect to  $\odot$ , or in other words,  $(x \odot y) \oplus z \neq (x \oplus z) \odot (y \oplus z)$ .*

*Proof.* The above properties can be easily derived from Equations 12 , 13 and 14.  $\square$

Note that the path  $P(x_1, x_n)$  joining  $x_1, x_2, \dots, x_n$  is equivalent to

$$\bigodot_{i=1}^n x_i. \quad (15)$$

We also note that if two paths  $P(x_1, x_n)$  and  $P(y_1, y_m)$  converge into the same node  $z$  as depicted in Figure 24, this can be written as

$$\left( \bigodot_{i=1}^n x_i \right) \oplus \left( \bigodot_{i=1}^n y_i \right) \rightarrow z. \quad (16)$$

#### 5.4. The Assessment of Potential Solutions

In this work, a solution is assumed to be based on the paths joining a set of nodes from a query and its neighbouring concepts. However, any such path needs to be assessed to determine whether it provides a viable solution. Loosely speaking, intuition focuses on finding *different paths* to a solution, which might lead to a better solution compared to a “conventional” one. In this section we will address the assessment and identification of the most suitable set of solutions via the *intuition* and *propagation indices*. In the rest of this section, for brevity we shall refer to a path  $P(x_1, x_n) = \bigodot_{i=1}^n x_i$  as a vector  $\underline{x}$ .

#### 5.4.1. Intuition Index

Recall that knowledge is associated with the union of three different networks

$$G = G_k \cup G_i \cup G_c \quad (17)$$

where

- $G_k$  is the (semantic) network associated with the existing knowledge within a specific setting,
- $G_i$  is the (semantic) network associated with the intuitive knowledge and
- $G_c$  is the (semantic) network associated with the contextualised knowledge.

We can assume that a conventional solution is embedded on  $G_k$ .

**Definition 4.** We define the innovation index between the nodes  $x_s$  and  $x_e$  following the path  $\underline{x}$ , as

$$i(\underline{x}) = \frac{|E_p(G \setminus G_k)(\underline{x})|}{|E_p(G)(\underline{x})|}, \quad (18)$$

where  $E_p(G \setminus G_k)(\underline{x})$  and  $E_p(G)(\underline{x})$  are the set of edges in  $G \setminus G_k$  and the set of edges in  $G$  for a path  $\underline{x}$  between the nodes  $x_s$  and  $x_e$ , respectively.

Hence, based on Equation 18, the overall innovation index between the nodes  $x_s$  and  $x_e$ , assuming there are  $n$  paths between them is

$$i(x_s, x_e) = \frac{1}{n} \sum_{j=1}^n i_{\underline{x}_j} \quad (19)$$

#### 5.4.2. Propagation Index

As discussed in Section 5.3.2, the information is propagated based on specific rules, such as Equation 12 and more specifically, Equation 13. We define the *propagation index* associated with a path  $\underline{x}$  as

$$\alpha(\underline{x}) = \prod_{x_i \in \underline{x}} \alpha(x_i). \quad (20)$$

The propagation index simply estimates how well information can spread along a specific path  $\underline{x}$  and it will be used in conjunction with the innovation index to assess the suitability of paths related to a solution.

#### 5.4.3. Edge Entropy

Entropy is the measure of the average information content one is missing when one does not know the value of the random variable (Shannon, 2001). More specifically, entropy is a measure of information disorder. Entropy plays an essential role in information theory, whose goal is to assess the level of relevant and accurate information that is shared across one or more systems. The concept of information entropy has various (yet similar) formulations, which aims to explicitly evaluate the overall information level associated with specific system configurations (Trovati et al., 2019). Loosely speaking, high entropy values correspond to ‘information disorder’ which refers to a scenario not suitable for reasoning.

A network with very strong edge relations is optimal for reasoning as information is propagated more reliably and as such, the corresponding scenario can be modelled more efficiently. Therefore, we can use the concept of entropy to explore and assess how well a network with respect to a specific query-concept can be used to reason and find a solution. Note that high entropy values are also associated with very sparse networks. However, this corresponds to a trivial case, which will not be discussed in this work.

**Definition 5.** Let  $\underline{x}_t$  be a path at time  $t$  (where an edge is added at each time iteration) from a given node  $x_s$ . We define its entropy relative to as  $\underline{x}_t$ .

$$H(\underline{x}_t) = -\alpha(\underline{x}_t) \log \alpha(\underline{x}_t). \quad (21)$$

The overall entropy from  $x_s$  is given by

$$H(x_s) = -\sum_{j=1}^k \alpha(\underline{x}_t^j) \log \alpha(\underline{x}_t^j), \quad (22)$$

for all the paths  $\underline{x}_t^1, \dots, \underline{x}_t^k$  originating from  $x_s$ .



The concept of entropy defined in Definition 5 will be used as an exploratory tool to assess whether a specific query (defined by one or more concepts) can be reasoned upon based on a specific knowledge network. In other words, this will allow us to identify:

- Whether a given network can be used to identify one or more solutions from a query;
- How far we need to navigate into the network (by following paths originating from the query/concepts) to obtain a feasible solution.

It is straightforward to prove the following lemma.

**Lemma 3.** *Let  $H(x_s)$  be defined as in Equation 22. If  $\alpha(\underline{x}_t^j) = 1/e$  for all the paths  $\underline{x}_t^1, \dots, \underline{x}_t^k$  originating from  $x_s$ , then  $H(x_s)$  is at its maximum value.*

The rest of the section will focus on defining a discovery algorithm to identify the best solution(s) for a query. The aim is to approximate the value of the entropy of the different paths, which are incrementally expanded during the discovery process. This will allow to automatically assess the most appropriate solution based on the edge properties of the corresponding paths.

As discussed above, it is clear that the best outcome in assessing a path  $\underline{x}$  is when all the edges are associated with an activation index equal, or very close to 1. Let  $0 \leq \epsilon \leq 1$  and consider the values of  $H(\underline{x})$  for  $\alpha(\underline{x}) = 1 - \epsilon$ .

**Proposition 2.** *Let  $0 \leq \epsilon \leq 1$ . We have that the entropy of the path  $\underline{x}_t$  at the time iteration  $k$  is*

$$\begin{aligned} H(x_s) &= \sum_{i=1}^k H(\underline{x}_t^i) \approx \sum_{i=1}^k \left( \epsilon_i - \frac{\epsilon_i^2}{2} - \frac{\epsilon_i^3}{6} - \frac{\epsilon_i^4}{12} - \frac{\epsilon_i^5}{20} \right) \\ &\leq k \left( \bar{\epsilon} - \frac{\bar{\epsilon}^2}{2} - \frac{\bar{\epsilon}^3}{6} - \frac{\bar{\epsilon}^4}{12} - \frac{\bar{\epsilon}^5}{20} \right) \text{ where } \bar{\epsilon} = \max_{i=1}^k \{\epsilon_i\}. \end{aligned} \quad (23)$$

*Proof.* The above can be obtained via Taylor's expansion of  $x \log x$  at  $x = 0.5$  and by substituting  $x = 1/2 - \epsilon$ , namely

$$\sum_{n \geq 2} \frac{(-1)^{1+n} 2^{-1+n} (-0.5 + \epsilon)^n}{(-1 + n)n} + (\epsilon - 0.5)(\log(2) - 1) + \frac{\log(2)}{2} \quad (24)$$

Therefore, for all the paths discovered at time  $t$ , which originate at the node  $x_s$ , the result follows from Equation 22 .  $\square$

Note that if  $\bar{\epsilon}$  is close to all the  $\epsilon_i$  and if  $\bar{\epsilon}$  is also close to 0, we say that the query has a *strong solution space* at time  $t$ .

During the exploration process, the network originating from a query concept is expanded by considering more connected nodes and edges. As discussed above, this process will end when the overall information (quantified by the edge entropy) is sufficient to describe a specific query space. Let  $E_q$  be the network associated with a query concept  $q$  and define a  $|E_q|$ -dimensional space

$$\mathcal{E}_q = [0, 1]^{|E_q|}. \quad (25)$$

Consider the path  $\underline{x}_t$  with weights  $\alpha_1, \dots, \alpha_k$  at the time iteration  $t$ . Let  $\mathcal{C}_{\underline{x}_t}$  be the hypercube in  $\mathcal{E}_q$  with sides having length  $\alpha_1, \dots, \alpha_k, \underbrace{1, \dots, 1}_{\text{for } k+1, \dots, |E_q|}$ .

In other words,  $\mathcal{C}_{\underline{x}_t}$  has sides equal to 1 if they do not correspond to any edge in  $\mathcal{E}_q$ .

### 5.5. Entropy of the Query Space

As discussed above, the concept of entropy captures the level of information disorder associated with a specific system. Entropy has been defined to quantify individual paths in a query space. However, a similar entropy definition can be used to assess the suitability of a (dynamical) query space with respect to a query. In other words, this can be used to evaluate a solution defined by the corresponding network. Loosely speaking, a high information disorder, which is associated with high levels of entropy, potentially refers to 'bad' decisions. In other words, the query space does not provide the

relevant information to explore the most suitable solutions based on a query concept.

The entropy defined in this section is solely based on the information propagation values associated with the different edges. In particular, this will justify the use of information propagation as the main tool to assess and explore the best solution(s) for a query. During this process, specific nodes may be chosen as a destination, which will be referred to as *leaf nodes*.

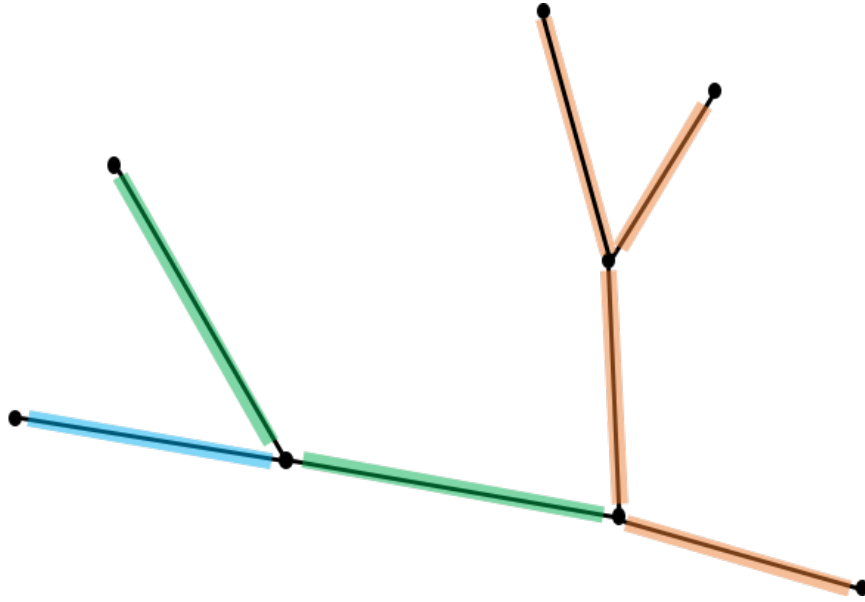


Figure 25: Different highlighted paths in a small network.

Based on Definition 5, the overall entropy can be used to explore the query space. From Lemma 3, we therefore need to consider the entropy in the interval  $[1/e, 1]$ . This is equivalent to having  $\alpha_k > 1/(ke)$

We will define the following algorithm:

Algorithm 4 return a path  $\underline{x}_t$  and its entropy based on a query. However, it does not guarantee to identify the shortest path between the concept and leaf nodes. Furthermore, based on the different parameters associated with the path,  $\underline{x}_t$  might not reach any of the leaf nodes. Despite such limitations, Algorithm 4 formalises an effective exploratory tool to assess a query. This will be further discussed in Section 6.

---

**Algorithm 4** Solution Assessment

---

```
1: Let  $t = 1$  and  $D \leq 1$  be the threshold for the maximum entropy
2: Let  $\underline{x}_t$  be a path connecting a query concept with one of its leaves  $\text{leaf}(\underline{x})$ 
3: for path  $\underline{x}$  do
4:   if  $\alpha(\underline{x}_t) \geq 1/(te)$  and  $H(\underline{x}_t) \leq D$  then
5:     Continue
6:   else
7:     Stop
8:   end if
9:   if  $\underline{x}_t \neq \text{leaf}(\underline{x})$  then
10:     $t = t + 1$ 
11:   else
12:     Stop
13:   end if
14: end for
15: return path  $\underline{x}_t, H(\underline{x}_t)$ 
```

---

### 5.5.1. Geometric Interpretation of the Query Space

Let

$$\prod_{i=1}^k \alpha_k = \text{Vol}(\mathcal{C}_{\underline{x}_t}), \quad (26)$$

where  $\text{Vol}(\mathcal{C}_{\underline{x}_t})$  is the volume of the corresponding hypercube. We then have the following result.

**Lemma 4.** *Let  $\mathcal{C}_{\underline{x}_t}$  be as defined above. Clearly we have that*

$$\text{Vol}(\mathcal{C}_{\underline{x}_0}) \geq \text{Vol}(\mathcal{C}_{\underline{x}_t}), \quad (27)$$

for  $t \geq 0$ .

In general, for  $t \leq T$ ,

$$\text{Vol}(\mathcal{C}_{\underline{x}_t}) \geq \text{Vol}(\mathcal{C}_{\underline{x}_T}). \quad (28)$$

*Proof.* For  $t \leq T$ ,  $\mathcal{C}_{\underline{x}_t}$  has more (or the same) default unit lengths compared to  $\mathcal{C}_{\underline{x}_T}$ . Therefore,  $\text{Vol}(\mathcal{C}_{\underline{x}_t}) \geq \text{Vol}(\mathcal{C}_{\underline{x}_T})$  for  $t \leq T$ . The results then follow.  $\square$

### *5.6. Summary*

In this chapter, a mathematical formulation and algorithm for artificial intuition was introduced. The mathematical formulation describe a model that utilises network to improve decision making. The model included some lemmas and propositions that provides a way of combining the aggregation of edges to discover new connections and pathways. The algorithm used the concept of entropy to explore the query space in order to identify potential solutions. These mathematical concepts and algorithms are evaluated in the next chapter.

## 6. Model Evaluation and Results

In this chapter, the evaluation of the method described in section 5 is presented. The evaluation is based on scenarios captured by semantic networks. To achieve the evaluation, a large semantic network was designed based on a suitable textual analysis of ConceptNet, Wikipedia and Recipe-Ingredient datasets. First, the knowledge-based will be introduced and discussed. More specifically, this is based on ConceptNet (Speer and Havasi, 2012), Wikipedia (McNeill, 1994) as a corpus and Recipe-Ingredient datasets available from Kaggle (Hoque et al., 2019) which are described in Sections 6.1.1, 6.1.2 and 6.1.3 respectively. The main motivation to use a knowledge-based approach is based on the simple observation that intuition is informed by any general knowledge, as well as more contextualised and ‘intuitive’ knowledge. The creation of a large network defined by such types of knowledge, namely ConceptNet, Wikipedia and Recipe-Ingredient respectively, is essential in designing an Artificial Intuition framework.

As described in Section 5.2, the study define an undirected network  $G = G(V, E)$ , where  $V = \{v_i\}_{i=1}^n$  is the *node set* and  $E = \{e_{w_{i,j}}(v_i, v_j)\}_{v_i \neq v_j \in V}$  is the *edge set*. Typically, each edge  $e_{w_{i,j}}(v_i, v_j)$  is associated with a weight  $w_{i,j}$ , which is related to the relationship between  $v_i$  and  $v_j$ . The study consider the network generated by the union of three (usually overlapping) different networks

$$G = G_k \cup G_i \cup G_c \quad (29)$$

where

- $G_k$  is the (semantic) network associated with the existing knowledge within a specific setting,
- $G_i$  is the (semantic) network associated with the intuitive knowledge and
- $G_c$  is the (semantic) network associated with the contextualised knowledge.

Using this formulation, we assume that ConceptNet is associated with  $G_i \cup G_c$ , and Wikipedia with  $G_k$ . Similarly, ConceptNet is associated with  $G_i \cup G_c$ , and Recipe-Ingredient with  $G_k$ .

### 6.1. *Description of Datasets: ConceptNet, Wikipedia and Recipe-Ingredient Datasets*

ConceptNet, Wikipedia and recipe-ingredient datasets have been widely used in respective researches. For example, ConceptNet has been widely used for various applications such as query answering (Kotov and Zhai, 2012); exploring Network Analysis of Knowledge Bases (Berger-Wolf et al., 2013); used in combination with other sources of distributional semantics (such as word2vec) to produce new embeddings with state-of-the-art performance across many word-relatedness evaluations (Mikolov et al., 2013; Pennington et al., 2014) as well as being used in combination with WordNet or DBPedia. These datasets were chosen because they provide suitable commonsense knowledge that are relevant for the evaluation of an intuition based model. The use of such commonsense knowledge and reasoning helps to constrain the operation of the algorithm that generates the network and provides meaningfulness of semantic network.

#### 6.1.1. *ConceptNet*

ConceptNet is a large semantic network of common sense knowledge. The dataset is analysed using Python to assess the semantic network, via specific patterns. Access to the ConceptNet database is provided through a web API using JavaScript Object Notation (JSON) textual data format. ConceptNet is a semantic network of knowledge that consist of assertions of common sense knowledge encompassing the spatial, physical, social, temporal, and psychological aspects of everyday life. The knowledge graph connects words and phrases of natural language with labeled edges. Its knowledge is collected from many sources that include expert created resources and open Mind Common Sense corpus (Liu and Singh, 2004), a crowd-sourced knowledge project. Version 5.7 of ConceptNet has been recently published and it is derived from several sources (Speer et al., 2017).

The network is designed to represent the general knowledge involved in understanding language, improving natural language applications by allowing the application to better understand the meanings behind the words people use. ConceptNet is similar to WordNet, however the advantage of ConceptNet over WordNet is its integrated Natural language Processing Engine (Liu and Singh, 2004). ConceptNet contains over 21 million edges and over 8 million nodes and access to the database is provided through a web API using JavaScript Object Notation (JSON) textual data format. Some of the concepts are involved in a handful of assertions and each comes with a positive or negative score, a weight that it assigns to that edge. The more positive the weight, the more likely that the assertion is true; a negative weight means we should conclude from these sources that the assertion is not true (Speer et al., 2017). Figure 26 depicts an example which shows a high-level view of common sense knowledge related to the following concepts: ‘*alarm clock; wake up; get to bed early; check email; drink coffee; yawn; eat breakfast*’.

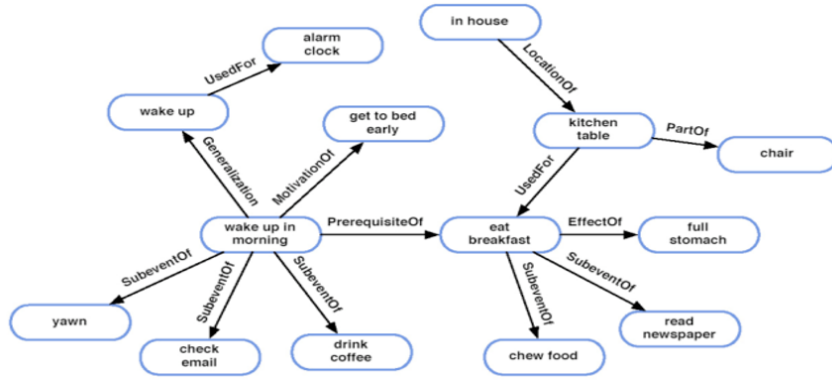


Figure 26: Example of some semantic relations from ConceptNet (Liu and Singh, 2004)

Similarly, Figure 27 shows an instance of the semantic network associated with the concept *car* including its related concepts. In this network, the notation  $IsA(Car, Vehicle)$  means that the concepts *Car* and *Vehicle* are connected by the directed relation *IsA*. Similarly,  $UsedFor(Vehicle, Mobility)$  means that the concepts *Vehicle* and *Mobility* are connected by the directed



relation *UsedFor* in this context.

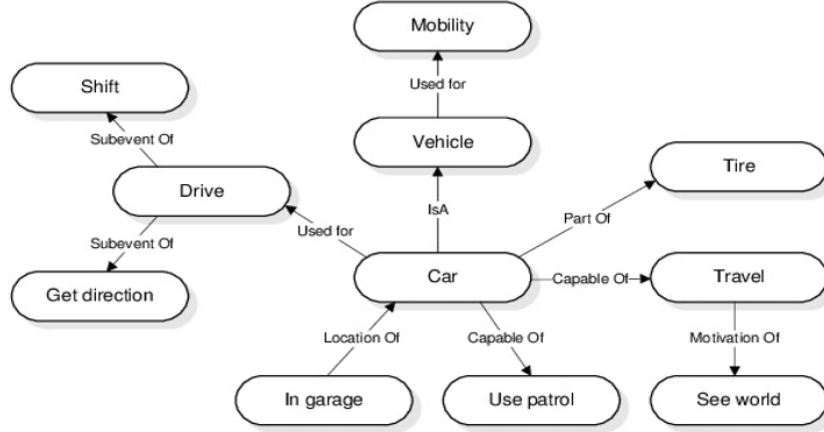


Figure 27: Semantic Network of a Car and Related Concepts (Liu and Singh, 2004)

In this study, ConceptNet is used to identify suitable semantic networks as part of the validation process, via the following steps: Details of the validation process is described in Sections 6.1.4

1. Identify specific concepts contained in a query;
2. Extract the relevant network defined by the main concepts related to the query and mutual relationships;
3. Build the semantic network of the main concepts and create a reduced semantic network in CSV
4. Merge the network with any other (semantic) network previously defined;
5. Navigate across the network to discover knowledge related to the query.

Access to the ConceptNet database is provided through a web API using JavaScript Object Notation (JSON) textual data format. However, due to performance reasons, we obtained the pre-built list of all the edges (assertions) in ConceptNet 5.7 and used ConceptNet-lite, a Python package to build it. This was used as an offline copy of ConceptNet 5.7 database. The offline copy was used because of the high volume of queries to the online ConceptNet during the creation of random semantic networks and

the application of variation operators. The offline copy of ConceptNet 5.7 database provides the complete dataset in locally accessible and highly efficient SQLite database format. This enables substantially faster access to data compared with the online version. Moreover, the offline copy contains more relations than the online version. Although there are some repeated assertions due to noise, the offline copy was found to be more useful for the evaluation of the model. The nodes of ConceptNet are words and phrases of natural language.

ConceptNet 5.5 has core set of 36 asymmetric and symmetric relations such as *is\_a*, *used\_for*, *part\_of*, *capable\_of*, *similar\_to*, *located\_near*, *related\_to* and these are intended to represent a relationship independently of the language or the source of the terms it connects.

An important decision to make when building a knowledge graph from ConceptNet, is on what a node should represent as this has significant effects on the graph that is retrieved and how the graph is used. Moreover, it also has implications that can make linking and importing other resources non-trivial, because different resources make different decisions about their representation (Speer and Havasi, 2012). Concepts are involved in a handful of assertions and each comes with a positive or negative score, a weight that it assigns to that edge. The more positive the weight, the more likely that the assertion is reliable and true; a negative weight means we can conclude from these sources that the assertion is not true (Speer and Havasi, 2012). A typical weight is 1.0, and the number is higher when the information comes from more sources or more reliable sources.

#### 6.1.2. *Wikipedia*

Wikipedia is a multilingual online encyclopaedia created and maintained as an open collaboration project by a community of volunteer editors using a wiki-based editing system (McNeill, 1994). It is the largest and most popular general reference work on the World Wide Web and it features exclusively free content. The data from Wikipedia is largely unstructured. An important requirement of the model described in this work is to represent knowledge as graph. Essentially, the challenge is to transform this text data

into something that can be used by the machines and also can be interpreted by us. To do this, spaCy (Choi et al., 2015) is used to build an appropriate graph representation of knowledge from Wikipedia pages. This was done in order to provide an additional database of structured information and relationships for knowledge discovery. This database is integrated and combined with ConceptNet to provide the appropriate knowledge discovery for the given scenarios modelled by the approach.

By using spaCy libraries, we were able to perform named entity recognition (NER). Natural Language Processing (NLP) such as sentence segmentation, Part of speech (POS) to extract pairs of entities and their relations from Wikipedia pages were applied in order to build the knowledge network. Infact, based on a specific query, the corresponding concepts are used to identify the relevant Wikipedia pages via the spaCy package, which creates a textual dataset. This was subsequently analysed with NLTK to identify relevant concepts and their mutual relationships based on the syntactic properties of the corresponding sentences.

#### *6.1.3. Recipe-Ingredient Dataset*

The recipe ingredients datasets is provided by <https://www.yummly.com/> and downloaded from Kaggle.com. The recipe dataset has featured in kaggle and has been used widely in Kaggle competition. The dataset contains different ingredients that cut across 3 continents (Europe, Latin America and Asia). The dataset includes the recipe id, type of cuisine and list of ingredients. Cuisine represents the country in which the ingredients occur in the recipes. The dataset is a list of 4088 unique ingredients and their corresponding recipes. Table 3 shows sample raw data of the recipe ingredient dataset.

The recipe ingredient data are stored in JSON format. The first step in processing the data is to convert the json file to csv format. Next we stored the data in dictionary and converted it to dataframe. We did the feature selection. As we are only interested in the recipe and ingredient data, we removed the recipe id. Moreover, a network representation of the data, called FoodNet was created. This was done in order to combine it with

ConceptNet to create the appropriate network knowledge representation for the evaluation of the model.

Country Cuisine	Ingredients
mexican	vegetable_oil
italian	mozzarella_cheese
spanish	diced_red_onions
mexican	garlic_salt
italian	large_shrimp
thai	curry_paste
mexican	onions
italian	olive_oil
mexican	garlic
italian	ground_black_pepper
mexican	chili_powder
italian	garlic
japanese	pepper
indian	tumeric

Table 3: Sample raw data of the recipe ingredient dataset

#### 6.1.4. Validation Approach

The validation approach is shown in Figure 28

Below is the details illustrations of the steps and components involved in the validation of the artificial intuition model

1. Load the respective datasets into dataframe
2. Create the semantic network for the respective datasets
3. Compare and combine the networks
4. Get all the paths between two nodes with a certain depth
5. Get all the edges (as node pairs) from a path
6. get the weights of a sequence of edges from each of the paths
7. combine the paths with the weight
8. Given two nodes, get all the paths and corresponding weights

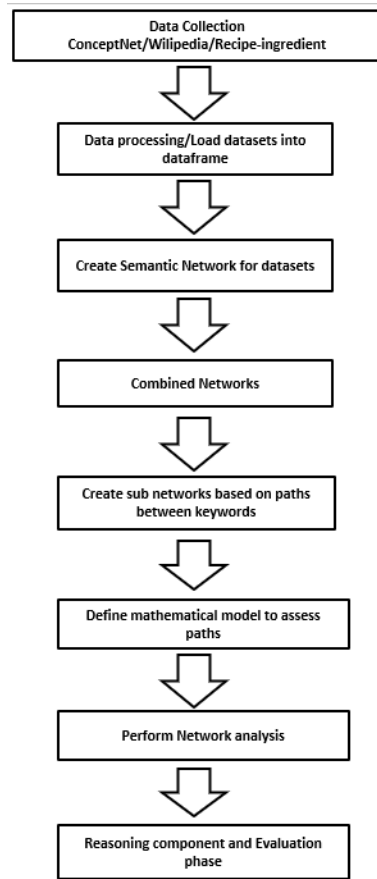


Figure 28: The components in developing the Artificial Intuition Model

9. Create the sub network based on the paths between two keywords (e.g between *isi\_ewu* and *Mexico*)
10. Create and differentiate two sub-networks (from FoodNet and conceptNet)
11. Define the mathematical model to assess the paths.
12. Perform the network analysis
13. Develop a reasoning strategy to navigate a path from *isi\_ewu* to *Mexico*

## 6.2. Evaluation and Results

The first part of the validation process presented in this study, focuses on a query ‘weather forecast’. Weather forecast modelling has been investigated by a large research community (Ghelli, 2015) as an accurate prediction requires the implementation of complex algorithms to address the most significant parameters. There are, however, emerging approaches based on new techniques to provide more agile and computationally efficient models (Hewage et al., 2020).

The assertions associated with the following concepts were retrieved: *weather, rain, rainfall, wind, temperature, cloud, cloudy, modelling, maths, statistics, sunshine, hot\_weather, weather\_forecast, weather\_prediction predict\_rainfall, wind\_forecast, maths\_statistics*.

Initially, over 5495 relations from ConceptNet database was retrieved. The lowest weight in the relations retrieved was 0.1 while the highest weight was 10.472. It was noticed that there were some repeating assertions which are due to noise. The issue of noise in the assertions is not surprising because the data are contributed by different sources. This issue of noise was addressed by ignoring all assertions with weight less than 1. The weights were subsequently normalised and grouped into 10 discrete categories  $[0, 0.1), [0.1, 0.2), \dots [0.9, 1]$ .

Although it was specified that only English data should be retrieved, however, some non-english assertions were noticed. Further data cleansing was done in order to prepare the resultant semantic network for analysis. We preprocessed the data and created the retrieved semantic network in a CSV database. This was done in order to use Pandas dataframe and NumPy on the CSV database and to perform some analysis of the network.

Another important consideration is to define the depth of each of the relation structures in our retrieved semantic network from ConceptNet. Because ConceptNet contains millions of edges and nodes, an important decision to make when building a knowledge graph from ConceptNet, is on what concepts to use and relations that are involved in these concepts. Note that an important characteristic of ConceptNet is that a given pair of concepts can be linked by multiple relation types, and relations can have multi-word

	Path length	Propagation Index	Innovation Index
Aircraft	9	0.077	0.72
Climate Change	5	0.39	0.63
Statistics	3	0.64	0.39
Cattle	11	0.20	0.45

Table 4: The validation results as discussed in Section 6.2

arguments of diverse semantic types. These places relations in close vicinity in semantic space, making relation prediction a hard task. On average this applies to 5.37% of instances per relation (Becker et al, 2019).

The relations in ConceptNet can be as deep as the representation requires within the semantic network. Because the size of the network can grow exponentially, parameters was used to limit the size of the network. In other words, we prune the resultant semantic network. Limiting the sizes of the relation structures reduces the noise and improves the performance and efficiency of the resultant network. Moreover, the reduced semantic network makes it easier and simpler to work with. Given the relation structures of each of the concepts in the dataset, we can define the depth of the relation structure. The more the dept the more the related concepts that are retrieved. In this experiment, relation grouping and weight importance is used. It was found that the weight importance of between 1 and 4 were reasonable choices in the evaluation using ConceptNet dataset. The query runs in approximately 30 seconds on reduced ConceptNet and in 60 minutes on the full ConceptNet.

This query identified the concepts: *aircraft*, *climate change*, *statistics* and *cattle*, as shown in Table 4. Note that despite a low propagation index for *aircraft*, its innovation index is the highest. Interestingly, aircraft technology does contribute to weather forecast, despite not been fully captured in the dataset created for this validation.

#### 6.2.1. Evaluation using Recipe-ingredient and ConceptNet Dataset

The second part of the validation process focuses on recipe-ingredient and ConceptNet dataset. For this we retrieved the assertions associated with the following concepts: *food*, *recipes*, *spices*, *spicy*, *ingredient*, *spicy\_food*.

The second part of the validation focused on food recipes related to West African delicacies, such as ‘*nkwobi*’, ‘*isi\_ewu*’ and ‘*suya*’. The aim was to find suitable solution for making these delicacies. To do this, ConceptNet was used despite no category related to the West African food is present. Via the model discussed in this study, suitable semantic connections among the ingredients for making ‘*nkwobi*’, ‘*isi\_ewu*’ and ‘*suya*’ were identified. More specifically, the focus was also on a query regarding South American food. Loosely speaking, this was equivalent to identifying a link between West African and South American recipes.

Subsequently, suitable combination of ingredients that can potentially improve specific cultural entity of Latin American cuisine were considered.

By using the model described in this study, we were able to integrate and combine the recipe ingredient network and ConceptNet to identify suitable semantic connections that are potentially capable of creating relevant recipes for the food. The following steps were implemented to achieve the result:

- A semantic network related to food concepts from the recipe ingredient dataset was created. This network is called FoodNet
- A semantic network related to food concepts from ConceptNet was created.
- Various occurrences of ingredients were identified from FoodNet and in the extracted semantic network database of food from ConceptNet.
- If two ingredients co-occur, regardless of the direction of the connecting edge, we say they are strongly linked.
- This was used to further identify food categories containing specific ingredients. Such categories were also enriched by querying the ConceptNet database to identify the corresponding synonym for each of the concepts.
- Each synonym relation found in ConceptNet was compared with WordNet synsets, to confirm the reliability of the synonymity of the relation that are used in the food connection/pairing. We found that all the



synonym relations with weight equal to 2 in ConceptNet were returned by WordNet.

- The three networks  $G_k, G_c$  and  $G_i$  as per Equation 4 were chosen manually from the dataset created for this validation.

We used various combination of ingredients in the analysis and was able to find links to ConceptNet data for making specific foods. For example, note that there was no category for African food in the FoodNet, however, by using the model described, we were able to find suitable semantic connections between FoodNet and ConceptNet for ingredients for making *nkwoobi*, *suya* and *Isi.ewu*. *nkwoobi*, *suya* and *Isi.ewu* are three West African cuisines that contains spicy ingredients. We found occurrences of *chili*, *ginger*, *pepper*, *garlic*, *spices* in FoodNet which was augmented with ConceptNet to discover the path to the solution.

After identifying the ingredient to make the specific African cuisines (*suya*, *nkwoobi*, *isi.ewu*), we further used our approach to find suitable combination of ingredients that can potentially improve specific cultural entity of Latin American cuisine. To do this, we constructed another Food network from the recipe ingredient dataset to capture the knowledge about ingredients and their country's cuisine in which they are used. This is a weighted undirected network and it consists of ingredients and country's cuisines as the nodes. The resulting entire network is too dense for visualisation. Therefore, we visualised subsets of the network. We used a cutoff of depth 5, that is any path greater than the depth of 5 is cutoff. We did this to reduce the density of the network. See Figure 29 for the dense ConceptNet network.

We then determined the weight of each edge by calculating the frequency of occurrence of the ingredients within the particular country's cuisines. Ingredients that frequently occurred in the network has a higher weight. We used a weight between 4 and 1 with 1 being the highest. The weight of the edge between each set of nodes represent the reliability and the probability of occurrence in the network. The average number of ingredients used per country's cuisine is seven. Some of the significant ingredients that occur together were identified. The application of the edge activation of the network

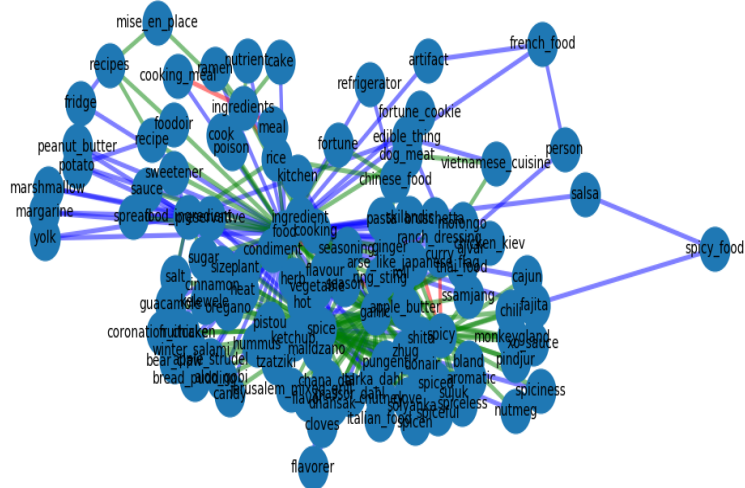


Figure 29: Graph of the weighted paths of ConceptNet

was illustrated by asking the question:

*"What can I use to improve the Latin American cuisine?"*

In this context, we assumed that this query cannot be directly addressed within the existing ingredient knowledge network. Therefore, suitable semantic links between the general knowledge and contextual knowledge will need to be assessed for the specific problem domain. This was achieved by exploring the resulting combined network and the edges joining the concepts related to the above query. The paths connecting these concepts were assessed to identify potential solutions. To achieve this, the FoodNet is

combined with ConceptNet to create a combined semantic network. Next, we performed the queries on the resultant combined network based on some keywords and conditions. More importantly, we identified the edge joining these queries and explored potential new solution. The query activated over 500 nodes using a cutoff of 5. The intermediate nodes that frequently occurred in the query are *spicy*, *garlic*, *chili*. Because the network as a whole could not be examined, related nodes in the network were identified and by using the contextual knowledge it was possible to find new links and discover the paths from an African cuisine to Latin American cuisine. This was done by performing analysis on contextual information and use that to augment the edges from the general knowledge. See the paths involving some of the activated pairs of nodes. Note that the first two pairs are from ConceptNet sub network while the second pair are from FoodNet.

- [*isi\_ewu*– > *spicy*– > *garlic*– > *mexican*]
- [*isi\_ewu*– > *spicy*– > *food*– > *mexican*]
- [*isi\_ewu*– > *spicy*– > *chili*– > *mexican*]

Starting with the node *isi\_ewu*, the relevant sub-network was identified. Each node activates new associative paths and by using the contextual knowledge it was possible to use the model to discover the paths leading to the solution. The edge weight were used to get the weighted paths between the nodes and prune the network in order to activate a more efficient solution. With the use of a query, a subnet was created using the paths containing relevant keywords (e.g. *isi\_ewu* and *mexican*) from the merged networks. See Figure 30 for the sub-networks between FoodNet and ConceptNet with depth of 4. Figures 31 and 32 shows the respective paths to Mexican and Brazilian cuisines.

The blue nodes *isi\_ewu* and *spicy* are the sub network from ConceptNet while the green nodes are the subnets from FoodNet. By traversing the path from *isi\_ewu* to *garlic* via the node *spicy*, it is possible to use our model to discover a new path linking *mexican*. More specifically, we were able to jump from *garlic* to the solution based on the model’s contextual



```

[[['isi_ewu', 'spicy', 3], ['spicy', 'food', 2], ['food', 'mexican', 2]], [['isi_ewu', 'spicy', 3], ['spicy', 'garlic', 3], ['garlic', 'mexican', 1]], [['isi_ewu', 'spicy', 3], ['spicy', 'ginger', 3], ['ginger', 'mexican', 2]], [['isi_ewu', 'spicy', 3], ['spicy', 'hot_sauce', 2], ['hot_sauce', 'mexican', 4]], [['isi_ewu', 'spicy', 3], ['spicy', 'chili', 3], ['chili', 'mexican', 3]], [['isi_ewu', 'spicy', 3], ['spicy', 'habanero', 3], ['habanero', 'mexican', 2]], [['isi_ewu', 'spicy', 3], ['spicy', 'nutmeg', 3], ['nutmeg', 'mexican', 2]], [['isi_ewu', 'spicy', 3], ['spicy', 'pepper_jack', 3], ['pepper_jack', 'mexican', 4]]]

```

Figure 31: Paths from isi\_ewu to Mexican Cuisine (Latin American)

```

[[['isi_ewu', 'spicy', 3], ['spicy', 'garlic', 3], ['garlic', 'brazilian', 1]], [['isi_ewu', 'spicy', 3], ['spicy', 'ginger', 3], ['ginger', 'brazilian', 1]], [['isi_ewu', 'spicy', 3], ['spicy', 'hot_sauce', 2], ['hot_sauce', 'brazilian', 1]], [['isi_ewu', 'spicy', 3], ['spicy', 'chili', 3], ['chili', 'brazilian', 1]]]

```

Figure 32: Paths from isi\_ewu to Brazilian Cuisine (Latin American)

distribution of the dataset. This was carried out to categorise the ingredient distribution based on their likelihood of being spicy. Based on Ahn et al. (2011), Western cuisines show a tendency to use ingredient pairs that share many flavour compounds, supporting the food pairing hypothesis. By contrast, East Asian cuisines tend to avoid compound sharing ingredients. The authors state that ingredients sharing flavour compounds are more likely to taste well together than ingredients that do not. Our study is not to disagree or support this hypothesis, but rather to use our model to find a suitable network combination of ingredients for improving the cuisine of a particular cultural entity that has properties of spicy. We explored the networks of related spicy ingredients across the cultural entities.

An important component relevant to find the solution is the co-occurrence of the concepts in the context of the domain. This required to create a sub network within the FoodNet, by using the co-occurrence information of the ingredients. For example, two ingredients share an edge if they occur together more than would be expected by chance. See table 5 for the co-occurrence of ingredients in FoodNet. Accordingly, a classification of the ingredients using a calculation of the ingredients co-occurrence was constructed. In the query on the ingredient network (FoodNet), let's assume there are two important concepts that can potentially contribute to the solution domain. For example, ('garlic', 'pepper') and ('pepper', 'seasoning').

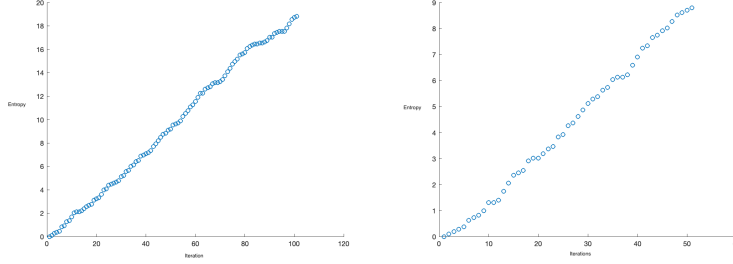


Figure 33: The values of the entropy of the query solution for the validations described in Section 6.2

If the weights of these relations in the network are low that reduces the strength of the connections, therefore further analysis has to be done in comparison with the other relation in the neighbourhood to find whether the strength is strong enough for it to be selected as part of the candidate terms or criteria to propagate or activate the solution. This processing goes on and on until the final candidate that can potentially improve the solution is selected. This is the light bulb moment. We found that our model has the capability to be adapted to recommend suitable combination of ingredients that can potentially improve the cuisine of some cultures.

Node1	Node2	Node1_count	Node2_count	CooC
garlic	pepper	7380	4438	1297
pepper	seasoning	4438	137	17
Black olives	garlic	229	7380	41
pepper	Black olives	4438	229	30
seasoning	Black olives	137	229	2
spices	garlic	171	7380	48
spices	pepper	171	4438	20
romaine lettuce	purple onion	270	1896	62
pepper	Feta cheese crumbles	4438	358	58

Table 5: Co-occurrence of ingredients in FoodNet

### *6.3. Summary*

In this chapter, the author have presented and discussed the evaluation of the model which utilises knowledge network based representation of the datasets. The evaluation is based on scenarios captured by a semantic network based on a suitable textual analysis of ConceptNet, Wikipedia and Recipe-Ingredient datasets. Some discussions about the validation approach was shown and finally the results of the evaluation was presented and discussed.

The author however, state that the validation of the model is not deep. In fact, this is far from being a comprehensive evaluation due to some obvious reasons. First of all, innovative and novel solutions can be complex to precisely pinpoint as this would require a manual comparison based on data that is currently either limited or non existent. Furthermore, ConceptNet and Wikipedia do not offer full semantic properties to create a fully implementable semantic network for artificial intuition. The purpose of the validation is not to provide a full implementation of the model, rather it is to demonstrate the value, soundness and validity of the approach when fully implemented. The full implementation of the model is outside the scope of this research and will be considered in future work.

## 7. Conclusion and Future Works

Despite artificial intuition has drawn considerable attention from the research community, there has been limited effort in defining and investigating its mathematical concepts and properties. This work aims to address this by introducing a rigorous approach to model and implement artificial intuition.

In this work, the author have discussed artificial intuition and decision making. The author provided a state-of-the-art review and provided a compelling motivation for the development of a computational model of artificial intuition and decision making. Relevant and detailed research about the concepts of artificial intuition as it relates to creativity, gut-feeling, rational thinking have been provided. The review finally identified the umbrella concept called artificial intuition and identified some key requirements for the development of a model. The author provided a rigorous modeling of artificial intuition in specific decision making scenarios. The author described the main mathematical formulations and algorithms that stems from the observation that intuition identifies new pathways between a starting set of concepts, to a specific target, which is the solution domain. The model focuses on finding different paths to a solution, which might lead to a better solution compared to a “conventional” one. In this approach, a solution is modelled by a path joining a set of nodes and such paths needs to be assessed to determine whether it provides a viable solution. The author posit that an approach that correctly implement the model can potentially obtain accurate and optimal result and improve the overall performance of human decision making systems.

The model was evaluated using a knowledge based network that is generated by the union of three different and overlapping networks. This network is based on scenarios captured by semantic networks that was designed based on a suitable textual analysis of ConceptNet, Wikipedia and Recipe-Ingredient datasets. The main motivation for using a knowledge-based approach is based on the simple observation that intuition is informed by any general knowledge, as well as more contextualised and ‘intuitive’ knowledge. The evaluation is provided to demonstrate the value, soundness and validity



of the approach when fully implemented. The preliminary experimental results demonstrate the potential of this approach and motivate further work in this field.

The main novelty and contributions of this work are:

1. The development of a sound mathematical formulation and algorithm for artificial intuition. Specifically, a mathematical formulation is introduced to describe a model that utilises semantic network to improve decision making. Moreover, the model that is being introduced included some lemmas and propositions that provides a way of combining the aggregation of edges. This can also lead to the possibility of creating the algebraic and proper mathematical theories to link the model with mathematical fields. This will potentially complement the computational model introduced in this research.
2. It implements techniques from computational linguistic via the processing pipeline to derive semantic networks. Facts, ideas and memories in our brain are not structured into tables like in relational databases, rather the brain is a graph, network of interconnected nodes. Therefore, using semantic network-based model and representation of the data is a suitable representation of how the brain works.
3. By using semantic network-based model, it is possible to apply spreading activation to simulate how human intuition work and quickly connect concepts thus improving the performance of intuitive decision making.
4. Contributed a state-of-the-art review of artificial intuition. The author provided relevant and detailed research about the concepts of artificial intuition as it relates to creativity, gut-feeling, rational thinking. It finally identified the umbrella concept called artificial intuition and identified some key requirements for the development of a model.

### *7.1. Limitation of the Research*

The main limitation of this research is that the validation of the approach is not deep. In fact, this is far from being a comprehensive evaluation due to

some obvious reasons. First, innovative and novel solutions can be complex to precisely pinpoint as this would require a manual comparison based on data that is currently either limited or non-existent. Second, there is no dataset ready and maintained for artificial intuition. In fact, ConceptNet and Wikipedia do not offer full semantic properties to create a fully implementable semantic network for artificial intuition. Moreover, the purpose of the evaluation presented is meant to demonstrate the value, soundness and validity of the approach. The validation presented was meant to show how the model will work when fully implemented. The full implementation will be considered in future work.

### *7.2. Future Work*

Future research will consider the full implementation of the model. In fact, future research will focus on a comprehensive mathematical analysis of the dynamical and algebraic properties of information aggregation and propagation. This will provide an in-depth analysis of critical aspects and properties which can be applied to artificial intuition and its integration with AI systems. Furthermore, a collaborative research effort will focus on the creation of a large semantic dataset (knowledge network) to complement and enhance the state-of-the-art data, which is currently available. This knowledge network will be specifically designed for an Artificial Intuition framework. It will enable a more comprehensive evaluation tools as well as training datasets to automatically identify some of the most important parameters related to artificial intuition. Moreover, human participants will be recruited to perform the implementation.

## References

- Aflague, J.M., Ferszt, G.G., 2010. Suicide assessment by psychiatric nurses: a phenomenographic study. *Issues in Mental Health Nursing* 31, 248–256.
- Ahn, H., 2010. Modeling and analysis of affective influences on human experience, prediction, decision making, and behavior. Ph.D. thesis. Massachusetts Institute of Technology.
- Ahn, Y.Y., Ahnert, S.E., Bagrow, J.P., Barabási, A.L., 2011. Flavor network and the principles of food pairing. *Scientific reports* 1, 196.
- Albert, R., Barabási, A.L., 2002. Statistical mechanics of complex networks. *Reviews of modern physics* 74, 47.
- Amirkhani, H., Rahmati, M., Lucas, P.J., Hommersom, A., 2016. Exploiting experts' knowledge for structure learning of bayesian networks. *IEEE transactions on pattern analysis and machine intelligence* 39, 2154–2170.
- Anderson, M., 2007. Artificial intuition: A new possible path to artificial intelligence. Recuperado de: <http://artificial-intuition.com/anderson.html>.
- Anderson, M.C., Bjork, R.A., Bjork, E.L., 1994. Remembering can cause forgetting: retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20, 1063.
- Arora, S., Barak, B., 2009. Computational complexity: a modern approach. Cambridge University Press.
- Aswath, D., Ahmed, S.T., D'cunha, J., Davulcu, H., 2005. Boosting item keyword search with spreading activation, in: The 2005 IEEE/WIC/ACM International Conference on Web Intelligence (WI'05), IEEE. pp. 704–707.
- Barnes, A., Thagard, P., 1996. Emotional decisions, in: Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society, University of California San Diego. pp. 426–429.

- Baronchelli, A., Ferrer-i Cancho, R., Pastor-Satorras, R., Chater, N., Christiansen, M.H., 2013. Networks in cognitive science. *Trends in cognitive sciences* 17, 348–360.
- Baydin, A.G., de Mántaras, R.L., Ontanón, S., 2015. A semantic network-based evolutionary algorithm for computational creativity. *Evolutionary Intelligence* 8, 3–21.
- Beer, R.D., 2000. Dynamical approaches to cognitive science. *Trends in cognitive sciences* 4, 91–99.
- Ben-Gal, I., 2008. Bayesian networks. *Encyclopedia of statistics in quality and reliability* 1.
- Berger-Wolf, T., Diochnos, D.I., London, A., Pluhár, A., Sloan, R.H., Turán, G., 2013. Commonsense knowledge bases and network analysis. *Commonsense*, May , Terjedelem–7.
- Berthold, M.R., Brandes, U., Kötter, T., Mader, M., Nagel, U., Thiel, K., 2009. Pure spreading activation is pointless, in: *Proceedings of the 18th ACM conference on Information and knowledge management*, pp. 1915–1918.
- Billay, D., Myrick, F., Luhanga, F., Yonge, O., 2007. A pragmatic view of intuitive knowledge in nursing practice, in: *Nursing Forum*, Wiley Online Library. pp. 147–155.
- Binali, H., Wu, C., Potdar, V., 2010. Computational approaches for emotion detection in text, in: *4th IEEE International Conference on Digital Ecosystems and Technologies*, IEEE. pp. 172–177.
- Blancke, S., De Smedt, J., 2013. Evolved to be irrational? evolutionary and cognitive foundations. *Philosophy of pseudoscience: Reconsidering the demarcation problem* , 341–379.
- Boden, M.A., et al., 2004. *The creative mind: Myths and mechanisms*. Psychology Press.

- Borge-Holthoefer, J., Arenas, A., 2010. Semantic networks: Structure and dynamics. *Entropy* 12, 1264–1302.
- Bower, G.H., Cohen, P.R., 2014. Emotional influences in memory and thinking: Data and theory. *Affect and cognition* 13, 291–331.
- Brachman, R.J., 1983. What is-a is and isn. *Computer* , 30–36.
- Brasethvik, T., Gulla, J.A., 2002. A conceptual modeling approach to semantic document retrieval, in: *International Conference on Advanced Information Systems Engineering*, Springer. pp. 167–182.
- Brien, S., Dibb, B., Burch, A., 2011. The use of intuition in homeopathic clinical decision making: an interpretative phenomenological study. *Evidence-Based Complementary and Alternative Medicine* 2011.
- Buckingham, C.D., Adams, A., 2000. Classifying clinical decision making: interpreting nursing intuition, heuristics and medical diagnosis. *Journal of Advanced Nursing* 32, 990–998.
- Budanitsky, A., Hirst, G., 2006. Evaluating wordnet-based measures of lexical semantic relatedness. *Computational linguistics* 32, 13–47.
- Ferrer-i Cancho, R., 2018. Optimization models of natural communication. *Journal of Quantitative Linguistics* 25, 207–237.
- Canese, K., Weis, S., 2013. Pubmed: the bibliographic database, in: *The NCBI Handbook* [Internet]. 2nd edition. National Center for Biotechnology Information (US).
- Carruthers, P., 2009. An architecture for dual reasoning. *J. St. BT Evans & K. Frankish* , 109–27.
- Cert, P.E., Wilcockson, J., 1996. Intuition and rational decision-making in professional thinking: a false dichotomy? *Journal of advanced nursing* 24, 667–673.

- Cheng, J., Bell, D., Liu, W., 1998. Learning bayesian networks from data: An efficient approach based on information theory. On World Wide Web at <http://www.cs.ualberta.ca/~jcheng/bnpc.htm> .
- Choi, J.D., Tetreault, J., Stent, A., 2015. It depends: Dependency parser comparison using a web-based evaluation tool, in: Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), Association for Computational Linguistics, Beijing, China. pp. 387–396. URL: <https://www.aclweb.org/anthology/P15-1038>, doi:10.3115/v1/P15-1038.
- Collins, A.M., Loftus, E.F., 1975. A spreading-activation theory of semantic processing. *Psychological review* 82, 407.
- Crestani, F., 1997. Application of spreading activation techniques in information retrieval. *Artificial Intelligence Review* 11, 453–482.
- Damasio, A.R., 1994. Descartes’ error: Emotion, rationality and the human brain .
- De Deyne, S., Navarro, D., Perfors, A., Storms, G., 2019. Structure at every scale: A semantic network account of the similarities between very unrelated concepts .
- De Deyne, S., Navarro, D.J., Perfors, A., Storms, G., 2016. Structure at every scale: A semantic network account of the similarities between unrelated concepts. *Journal of Experimental Psychology: General* 145, 1228.
- De Melo, C.M., Carnevale, P., Read, S., Antos, D., Gratch, J., 2012. Bayesian model of the social effects of emotion in decision-making in multiagent systems, in: Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 1, pp. 55–62.
- De Smedt, T., 2013. Modeling Creativity: Case Studies in Python. University Press Antwerp.

- Diaz-Hernandez, O., Gonzalez-Villela, V.J., 2017a. Analysis of human intuition towards artificial intuition synthesis for robotics. *Mechatronics and Applications: An International Journal (MECHATROJ)* 1.
- Diaz-Hernandez, O., Gonzalez-Villela, V.J., 2017b. Analysis of human intuition towards artificial intuition synthesis for robotics. *Mechatronics and Applications: An International Journal (MECHATROJ)* 1.
- Dreisbach, G., 2006. How positive affect modulates cognitive control: The costs and benefits of reduced maintenance capability. *Brain and cognition* 60, 11–19.
- Dreyfus, H., 1972. What computers can't do: The limits of artificial intelligence .
- Dundas, J., Chik, D., 2011. Implementing human-like intuition mechanism in artificial intelligence. *arXiv preprint arXiv:1106.5917* .
- Fellbaum, C., 2017. Wordnet: An electronic lexical resource, in: *The Oxford Handbook of Cognitive Science*. Routledge, pp. 301–314.
- Fenton-O'Creevy, M., Soane, E., Nicholson, N., Willman, P., 2011. Thinking, feeling and deciding: The influence of emotions on the decision making and performance of traders. *Journal of Organizational Behavior* 32, 1044–1061.
- Fenz, S., Tjoa, A.M., Hudec, M., 2009. Ontology-based generation of bayesian networks, in: *2009 International Conference on Complex, Intelligent and Software Intensive Systems*, IEEE. pp. 712–717.
- Friedman, N., Goldszmidt, M., 1998. Learning bayesian networks with local structure, in: *Learning in graphical models*. Springer, pp. 421–459.
- Gasper, K., 2003. When necessity is the mother of invention: Mood and problem solving. *Journal of Experimental Social Psychology* 39, 248–262.
- Ghelli, A., 2015. Visualization of confidence in weather forecasts: Ecmwf's users URL: <https://www.ecmwf.int/node/13315>.

- Gilovich, T., Griffin, D., Kahneman, D., 2002. *Heuristics and biases: The psychology of intuitive judgment*. Cambridge university press.
- Gmytrasiewicz, P.J., Lisetti, C.L., 2001. Emotions and personality in agent design and modeling, in: *International Workshop on Agent Theories, Architectures, and Languages*, Springer. pp. 21–31.
- Graber, M.L., Kissam, S., Payne, V.L., Meyer, A.N., Sorensen, A., Lenfestey, N., Tant, E., Henriksen, K., LaBresh, K., Singh, H., 2012. Cognitive interventions to reduce diagnostic error: a narrative review. *BMJ Qual Saf* 21, 535–557.
- Gratch, J., Marsella, S., 2005. Evaluating a computational model of emotion. *Autonomous Agents and Multi-Agent Systems* 11, 23–43.
- Halmos, F.G.P., 1982. *Graduate texts in mathematics* 84 .
- Hams, S.P., 2000. A gut feeling? intuition and critical care nursing. *Intensive and Critical Care Nursing* 16, 310–318.
- Hart, T., 2009. *From Information to Transformation: Education for the Evolution of Consciousness*. Counterpoints (New York, N.Y.), Peter Lang. URL: <https://books.google.co.uk/books?id=WtNWPgAACAAJ>.
- He, J., He, P., 2008. Fuzzy relationship mapping and intuition inversion: a computer intuition inference model, in: *2008 International Conference on MultiMedia and Information Technology*, IEEE. pp. 298–301.
- Hensman, A., Sadler-Smith, E., 2011. Intuitive decision making in banking and finance. *European Management Journal* 29, 51–66.
- Hewage, P., Behera, A., Trovati, M., Pereira, E., Ghahremani, M., Palmieri, F., Liu, Y., 2020. Temporal convolutional neural (tcn) network for an effective weather forecasting using time-series data from the local weather station. *Soft Computing* doi:10.1007/s00500-020-04954-0.
- Hoque, M.N., Coelho, D., Mueller, K., 2019. Examining the visualization practices of data scientists on kaggle. *IEEE* .



- Hudlicka, E., 2007. Reasons for emotions: modelling emotions in integrated cognitive systems. *Integrated Models of Cognitive Systems* 59, 1–37.
- Humphries, M.D., Gurney, K., 2008. Network ‘small-world-ness’: a quantitative method for determining canonical network equivalence. *PloS one* 3, e0002051.
- Isenberg, D.J., 1991. How senior managers think. Open University Press.
- Johnny, O., Trovati, M., Ray, J., 2017. The “gut-feeling” in the decision-making process: A computationally efficient approach to influence relation assessment, in: *International Conference on Intelligent Networking and Collaborative Systems*, Springer. pp. 576–584.
- Johnny, O., Trovati, M., Ray, J., 2019. Towards a computational model of artificial intuition and decision making, in: *International Conference on Intelligent Networking and Collaborative Systems*, Springer. pp. 463–472.
- Jones, M.N., Hills, T.T., Todd, P.M., 2015. Hidden processes in structural representations: A reply to abbott, austerweil, and griffiths (2015). .
- Kahneman, D., 2002. Maps of bounded rationality: A perspective on intuitive judgment and choice. *Nobel prize lecture* 8, 351–401.
- Kelly, J., Papalambros, P.Y., Seifert, C.M., 2008. Interactive genetic algorithms for use as creativity enhancement tools., in: *AAAI spring symposium: creative intelligent systems*, pp. 34–39.
- Kenett, Y.N., Levy, O., Kenett, D.Y., Stanley, H.E., Faust, M., Havlin, S., 2018. Flexibility of thought in high creative individuals represented by percolation analysis. *Proceedings of the National Academy of Sciences* 115, 867–872.
- Klein, G.A., 2017. Sources of power: How people make decisions. MIT press.
- Koestler, A., 1964. The act of creation. London Hutchinson.

- Kolb, P., 2009. Experiments on the difference between semantic similarity and relatedness, in: Proceedings of the 17th Nordic Conference of Computational Linguistics (NODALIDA 2009), pp. 81–88.
- Kotov, A., Zhai, C., 2012. Tapping into knowledge base for concept feedback: leveraging conceptnet to improve search results for difficult queries, in: Proceedings of the fifth ACM international conference on Web search and data mining, pp. 403–412.
- Kumar, G.R., Kumar, R.K., Sanyal, G., 2017. Facial emotion analysis using deep convolution neural network, in: 2017 International Conference on Signal Processing and Communication (ICSPPC), IEEE. pp. 369–374.
- La Pira, F., 2011. Entrepreneurial intuition, an empirical approach. *Journal of Management and Marketing Research* 6, 1.
- Lakoff, G., Johnson, M., 2008. *Metaphors we live by*. University of Chicago press.
- Langan-Fox, J., Shirley, D.A., 2003. The nature and measurement of intuition: Cognitive and behavioral interests, personality, and experiences. *Creativity Research Journal* 15, 207–222.
- Langley, P., Laird, J.E., Rogers, S., 2009. Cognitive architectures: Research issues and challenges. *Cognitive Systems Research* 10, 141–160.
- Leicht, E.A., Holme, P., Newman, M.E., 2006. Vertex similarity in networks. *Physical Review E* 73, 026120.
- Lerner, J.S., Li, Y., Valdesolo, P., Kassam, K.S., 2015. Emotion and decision making. *Annual review of psychology* 66.
- Lieberman, M.D., 2000. Intuition: a social cognitive neuroscience approach. *Psychological bulletin* 126, 109.
- Lipshitz, R., Shulimovitz, N., 2007. Intuition and emotion in bank loan officers’ credit decisions. *Journal of Cognitive Engineering and Decision Making* 1, 212–233.

- Liu, H., Singh, P., 2004. Conceptnet—a practical commonsense reasoning tool-kit. *BT technology journal* 22, 211–226.
- Liu, S., He, P., 2019. Artificial intuition reasoning system (AIRS) and application in criminal investigations. *Journal of Physics: Conference Series* 1302, 032032. doi:10.1088/1742-6596/1302/3/032032.
- Livet, P., 2010. Rational choice, neuroeconomy and mixed emotions. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365, 259–269.
- Loewenstein, G., Lerner, J.S., 2003. The role of affect in decision making. *Handbook of affective science* 619, 3.
- Lucas, P.J., Van der Gaag, L.C., Abu-Hanna, A., 2004. Bayesian networks in biomedicine and health-care. *Artificial intelligence in medicine* 30, 201–214.
- Lyneham, J., Parkinson, C., Denholm, C., 2008. Explicating benner’s concept of expert practice: intuition in emergency nursing. *Journal of advanced nursing* 64, 380–387.
- Madni, A., Moini, A., Madni, C., 2006. Cognitector™: Cognition-based, integrative architecture for next generation intelligent agents. Intelligent Systems Technology, Inc. Phase II Final Technical Report, ISTI-FTR-567-8/06, Contract .
- Marinier III, R.P., Laird, J.E., Lewis, R.L., 2009. A computational unification of cognitive behavior and emotion. *Cognitive Systems Research* 10, 48–69.
- Marra, R.M., Jonassen, D.H., 2002. Transfer effects of semantic networks on expert systems: Mindtools at work. *Journal of Educational Computing Research* 26, 1–23.
- Mayer, E.A., 2011. Gut feelings: the emerging biology of gut–brain communication. *Nature Reviews Neuroscience* 12, 453–466.

- McNeill, A., 1994. A corpus of learner errors: making the most of a database. Flowerdew and A. Tong (Eds.) , 114–125.
- Mikolov, T., Chen, K., Corrado, G., Dean, J., 2013. Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781 .
- Milgram, S., 1967. The small world problem. *Psychology today* 2, 60–67.
- Miller, C.C., Ireland, R.D., 2005. Intuition in strategic decision making: friend or foe in the fast-paced 21st century? *Academy of Management Perspectives* 19, 19–30.
- Minsky, M., 2007. The emotion machine: Commonsense thinking, artificial intelligence, and the future of the human mind. Simon and Schuster.
- Mousavi, S., Gigerenzer, G., 2014. Risk, uncertainty, and heuristics. *Journal of Business Research* 67, 1671–1678.
- Mueller, E.T., 2014. Commonsense reasoning: an event calculus based approach. Morgan Kaufmann.
- Murphy, R.R., Lisetti, C.L., Tardif, R., Irish, L., Gage, A., 2002. Emotion-based control of cooperating heterogeneous mobile robots. *IEEE transactions on robotics and automation* 18, 744–757.
- Nadkarni, S., Shenoy, P.P., 2004. A causal mapping approach to constructing bayesian networks. *Decision support systems* 38, 259–281.
- Naqvi, N., Shiv, B., Bechara, A., 2006. The role of emotion in decision making: A cognitive neuroscience perspective. *Current directions in psychological science* 15, 260–264.
- Navigli, R., Ponzetto, S.P., 2012. Babelnet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artificial Intelligence* 193, 217–250.
- Newman, M., 2018. *Networks*. Oxford university press.

- Ooi, B.C., Tan, K.L., Wang, S., Wang, W., Cai, Q., Chen, G., Gao, J., Luo, Z., Tung, A.K., Wang, Y., et al., 2015. Singa: A distributed deep learning platform, in: Proceedings of the 23rd ACM international conference on Multimedia, pp. 685–688.
- Oviedo, B., Moreira, L., Puris, A., Novoa, P., Moral, S., 2016. Learning bayesian network by a mesh of points, in: 2016 IEEE Congress on Evolutionary Computation (CEC), IEEE. pp. 3983–3989.
- Payne, L.K., et al., 2015. Intuitive decision making as the culmination of continuing education: A theoretical framework. *The Journal of Continuing Education in Nursing* 46, 326–332.
- Pease, A., Corneli, J., 2018. Evaluation of creativity, in: *Concept Invention*. Springer, pp. 277–294.
- Pennington, J., Socher, R., Manning, C.D., 2014. Glove: Global vectors for word representation, in: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp. 1532–1543.
- Pons, P., Latapy, M., 2005. Computing communities in large networks using random walks, in: *International symposium on computer and information sciences*, Springer. pp. 284–293.
- Poria, S., Cambria, E., Winterstein, G., Huang, G.B., 2014. Sentic patterns: Dependency-based rules for concept-level sentiment analysis. *Knowledge-Based Systems* 69, 45–63.
- Rada, R., Mili, H., Bicknell, E., Blettner, M., 1989. Development and application of a metric on semantic nets. *IEEE transactions on systems, man, and cybernetics* 19, 17–30.
- Rauterberg, M., 2010. Emotions: the voice of the unconscious, in: *International Conference on Entertainment Computing*, Springer. pp. 205–215.
- Redelmeier, D.A., Ferris, L.E., Tu, J.V., Hux, J.E., Schull, M.J., 2001. Problems for clinical judgement: introducing cognitive psychology as one more basic science. *Cmaj* 164, 358–360.

- Rocha, C., Schwabe, D., Aragao, M.P., 2004. A hybrid approach for searching in the semantic web, in: Proceedings of the 13th international conference on World Wide Web, pp. 374–383.
- Sadler-Smith, E., Shefy, E., 2004. The intuitive executive: Understanding and applying ‘gut feel’ in decision-making. *Academy of Management Perspectives* 18, 76–91.
- Sahlgren, M., Karlgren, J., 2008. Buzz monitoring in word space, in: European Conference on Intelligence and Security Informatics, Springer. pp. 73–84.
- Salichs Sánchez-Caballero, M., Malfaz Vázquez, M.Á., 2012. A new approach to modelling emotions and their use on a decision-making system for artificial agent .
- Schumacher, K., Sintek, M., Sauermann, L., 2008. Combining fact and document retrieval with spreading activation for semantic desktop search, in: European Semantic Web Conference, Springer. pp. 569–583.
- Segaran, T., Evans, C., Taylor, J., 2009. Programming the Semantic Web: Build Flexible Applications with Graph Data. ” O’Reilly Media, Inc.”.
- Sequeira, P., Melo, F.S., Paiva, A., 2016. Emergence of emotional appraisal signals in reinforcement learning agents, in: Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems, pp. 466–467.
- Shachter, R.D., Bhattacharjya, D., 2010. Solving influence diagrams: Exact algorithms. *Wiley Encyclopedia of Operations Research and Management Science* .
- Shannon, C.E., 2001. A mathematical theory of communication. *ACM SIGMOBILE mobile computing and communications review* 5, 3–55.
- Shastri, L., 1988. A connectionist approach to knowledge representation and limited inference. *Cognitive Science* 12, 331–392.

- Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., et al., 2016. Mastering the game of go with deep neural networks and tree search. *nature* 529, 484–489.
- Simon, H., Frantz, R., 2003. Artificial intelligence as a framework for understanding intuition. *Journal of Economic Psychology* 24, 265–277.
- Simon, H.A., 1987. Making management decisions: The role of intuition and emotion. *Academy of Management Perspectives* 1, 57–64.
- Sloman, A., 2001. Varieties of affect and the cogaff architecture schema, in: *Proceedings of the AISB’01 symposium on emotions, cognition, and affective computing*. The Society for the Study of Artificial Intelligence and the Simulation of Behaviour.
- Slovic, P., Finucane, M.L., Peters, E., MacGregor, D.G., 2007. The affect heuristic. *European journal of operational research* 177, 1333–1352.
- Smith, E.R., DeCoster, J., 2000. Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and social psychology review* 4, 108–131.
- Sonntag, D., 2011. Intuition as instinctive dialogue, in: *Computing with instinct*. Springer, pp. 82–106.
- Sowa, J.F., 1999. *Knowledge representation: logical, philosophical and computational foundations*. Brooks/Cole Publishing Co.
- Speer, R., Chin, J., Havasi, C., 2017. Conceptnet 5.5: An open multilingual graph of general knowledge, in: *Thirty-First AAAI Conference on Artificial Intelligence*.
- Speer, R., Havasi, C., 2012. Representing general relational knowledge in ConceptNet 5, in: *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC’12)*, European Language Resources Association (ELRA), Istanbul, Turkey. pp. 3679–3686.

URL: [http://www.lrec-conf.org/proceedings/lrec2012/pdf/1072\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2012/pdf/1072_Paper.pdf).

- Srdanov, A., Kovačević, N.R., Vasić, S., Milovanović, D., 2016. Emulation of artificial intuition using random choice and logic, in: 2016 13th Symposium on Neural Networks and Applications (NEUREL), IEEE. pp. 1–6.
- Srdanov, A., Milovanović, D., Vasić, S., Ratković Kovačević, N., 2017. The application of simulated intuition in minimizing the number of moves in guessing the series of imagined objects.
- Stanovich, K.E., West, R.F., 2000. Individual differences in reasoning: Implications for the rationality debate? *Behavioral and brain sciences* 23, 645–665.
- Suryadi, D., Gmytrasiewicz, P.J., 1999. Learning models of other agents using influence diagrams, in: *UM99 User Modeling*. Springer, pp. 223–232.
- Tanner, C.A., 2006. Thinking like a nurse: A research-based model of clinical judgment in nursing. *Journal of nursing education* 45, 204–211.
- Tao, W., He, P., 2009. Intuitive learning and artificial intuition networks, in: 2009 Second International Conference on Education Technology and Training, IEEE. pp. 297–300.
- Thiel, K., Berthold, M.R., 2010. Node similarities from spreading activation, in: 2010 IEEE International Conference on Data Mining, IEEE. pp. 1085–1090.
- Thompson, C., Cullum, N., McCaughan, D., Sheldon, T., Raynor, P., 2004. Nurses, information use, and clinical decision making—the real world potential for evidence-based decisions in nursing. *Evidence-based nursing* 7, 68–72.
- Tranel, D., Damasio, H., 1991. Somatic markers and the guidance of behaviour: Theory and preliminary testing.



- Trovati, M., Bessis, N., 2016. An influence assessment method based on co-occurrence for topologically reduced big data sets. *Soft computing* 20, 2021–2030.
- Trovati, M., Bessis, N., Huber, A., Zelenkauskaitė, A., Asimakopoulou, E., 2014. Extraction, identification, and ranking of network structures from data sets, in: 2014 Eighth International Conference on Complex, Intelligent and Software Intensive Systems, IEEE. pp. 331–337.
- Trovati, M., Castiglione, A., Bessis, N., Hill, R., 2015. A kuramoto model based approach to extract and assess influence relations, in: International Symposium on Computational Intelligence and Intelligent Systems, Springer. pp. 464–473.
- Trovati, M., Zhang, H., Ray, J., Xu, X., 2019. An entropy-based approach to real-time information extraction for industry 4.0. *IEEE Transactions on Industrial Informatics* 16, 6033–6041.
- Turney, P.D., 2008. A uniform approach to analogies, synonyms, antonyms, and associations. *arXiv preprint arXiv:0809.0124*.
- Tversky, A., 1977. Features of similarity. *Psychological review* 84, 327.
- Velásquez, J., 1998. Modeling emotion-based decision-making. *Emotional and intelligent: The tangled knot of cognition*, 164–169.
- Wanto, A., Windarto, A.P., Hartama, D., Parlina, I., 2017. Use of binary sigmoid function and linear identity in artificial neural networks for forecasting population density. *IJISTECH (International Journal of Information System & Technology)* 1, 43–54.
- Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of ‘small-world’ networks. *nature* 393, 440–442.
- Wilson, T.D., Schooler, J.W., 1991. Thinking too much: introspection can reduce the quality of preferences and decisions. *Journal of personality and social psychology* 60, 181.

- Wulff, D.U., Hills, T., Mata, R., 2018. Structural differences in the semantic networks of younger and older adults .
- Xu, Z., Srihari, S.N., 2014. Bayesian network structure learning using causality, in: 2014 22nd International Conference on Pattern Recognition, IEEE. pp. 3546–3551.
- Zeelenberg, M., Nelissen, R.M., Breugelmans, S.M., Pieters, R., 2008. On emotion specificity in decision making: Why feeling is for doing. *Judgment and Decision making* 3, 18.
- Zhang, Z., Zhong, W., 2016. Barriers to organizational creativity in chinese companies. Lewin, A., Kenney, M. and J. Murmann (eds.) *China's Innovation Challenge, Overcoming the Middle-Income Trap*. Cambridge University Press, Cambridge .